The American Journal of Human Genetics

Vol. 11 No. 4 *December 1959*

Some Comments on the Statistical Methods used in Linkage Investigations

CEDRIC A. B. SMITH

Galton Laboratory, University College London

NEWTON MORTON has shown in a series of papers (Morton 1955, 1957, Steinberg and Morton 1956) that sequential methods for the detection of linkage have many advantages. They are appreciably simpler and more efficient than most of the methods previously used, and are convenient in practice. Nevertheless their use raises important questions of principle. It seems to me that the use of a Wald sequential stop-rule, as advocated by Morton, is not appropriate in linkage work and confers no advantage: on the other hand Bayes' Theorem can be quite easily applied, using Morton's own results, and gives a more satisfactory answer from both the theoretical and practical points of view. (It may not in fact give a very different answer to that given by the sequential test: all the same it seems worth discussing the issues of principle involved.)

There are several lines of approach to the problems of statistical inference, namely (1) Bayes' Theorem, or inverse probability, (2) significance tests, (3) decision functions (4) sequential tests, (5) point estimation, including maximum likelihood, (6) interval estimation, including confidence and fiducial intervals. As these will be rather better known to the statistician than to the geneticist, it seems worth discussing the first four in some detail (for the others, see Kendall 1946).

(1) BAYES' THEOREM

I begin with a specific example. Consider a random mating population in which there are a pair of alleles *G*, *g*, each of frequency ½, and with *g* recessive to *G*. By the Hardy-Weinberg rule the genotypes will occur in the ratio ¼ *GG*:½ *Gg*:¼ *gg*. The intercross mating *Gg* × *Gg* will accordingly occur with frequency ½ × ½ = ¼, i.e. one mating in four will be of this type. The backcross mating *Gg* × *gg* can be either *Gg* ♀ × *gg* ♂ or reciprocally *Gg* ♂ × *gg* ♀, each with frequency ½ × ¼ = ⅛; to simplify the argument we will not distinguish between reciprocal matings, and so the backcrosses will form a proportion 2 × ⅛ = ¼ of all matings.

Suppose now that we have a sibship of tested children, say, for definiteness, one dominant (G) younger sib and one recessive (g = *gg*) older sib, but with untested parents. We wish to say as much as possible about the mating type of the parents. We see at once that it must be either backcross or intercross, since other matings produce all dominant or all recessive children. The question which remains is accordingly to find the relative chances of it being of one type or the other.

It is clear that we need only consider families of two children, like the one observed: other sizes of family will be irrelevant. Among all families of two, ¼ will be

Received November 29, 1958.

intercrosses and ¼ backcrosses. Since intercrosses produce (in the long run) ¾ dominant (G) children, and ¼ recessive (g), the proportion of intercrosses which produce one younger G and one older g type sib will be ¾ × ¼: thus, out of all families of two, a fraction ¼ × ¾ × ¼ = 3/64 will be intercrosses producing a younger type G sib and an older g: call these families of type T_1. A similar argument shows that a fraction

$$\frac{1}{4} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{16} = \frac{4}{64}$$

will be backcrosses producing a younger G and an older g, or, say, families of type T_2. Now apart from the phenotypes of the sibs we know nothing more about the observed family: we can accordingly imagine it as effectively chosen at random from among the whole group of families of types T_1 and T_2. Since the frequencies of these families are in the ratio $\frac{3}{64}T_1:\frac{4}{64}T_2$, i.e., three T_1 to every four T_2, it is evident that the observed family has chances in the ratio 3/64:4/64 = 3:4 of being type T_1 or T_2 respectively. The absolute probabilities of being T_1 and T_2 must accordingly be 3/7 and 4/7 respectively, since these are in the ratio 3:4 and sum to 1. Note that these can be calculated as

$$\frac{\frac{3}{64}}{\frac{3}{64} + \frac{4}{64}} = \frac{3}{7} \quad \text{and} \quad \frac{\frac{4}{64}}{\frac{3}{64} + \frac{4}{64}} = \frac{4}{7}$$

These probabilities have a quite definite meaning. If we select from the general population at random (or in an unbiased manner) all families with one younger dominant sib, and one older recessive, then a proportion 3/7 of them will be intercrosses, and a proportion 4/7 backcrosses. This is often called "inverse" probability, as it argues from an observed event to a previous one which has not been directly observed, or, as may be said, from "effect" to "cause." The general principle is as follows (Bayes 1763). Suppose that an event E can happen in k distinct ways W_1, W_2, ... W_k) (say, k possible "causes" of E). Suppose further that the way or cause W_r has a probability w_r of occurring in the general population, and that when W_r is known to have occurred, the event E has a probability e_r of following it. Then if we have observed that E has happened, but have no further knowledge about which way it happened, the probability that it did in fact happen through way W_r is

$$\text{Prob}\,(W_r|E) = w_r e_r/(w_1e_1 + w_2e_2 + \ldots + w_ke_k). \qquad (1)$$

The argument is essentially the same as in the genetical example, where the two "ways" W_1 and W_2 were the intercross and backcross mating respectively, with respective "initial" or "prior" probabilities w_1 = ¼, w_2 = ¼. The subsequent probabilities of getting the event E, i.e., the observed sibship, were e_1 = 3/16 from an intercross and e_2 = ¼ from a backcross. A substitution in (1) gives Prob $(W_1|E)$ = 3/7 as before.

To use formula (1) it is essential to know the initial probabilities w_r. This is made clear from a comparison of the following two situations:

(a) Matings of type G × g fall into two classes: either they are $GG \times gg$, and

produce only G children, or they are $Gg \times gg$, and produce equal numbers of G and g children. Suppose that a G × g mating taken at random gives rise to 10 dominant offspring: what is the chance that it is $GG \times gg$? If, as before, the gene frequencies of G and g are each $\frac{1}{2}$, the heterozygotes Gg are twice as frequent as the homozygotes GG, so the initial probabilities of the mating being $GG \times gg$ and $Gg \times gg$ are respectively $w_1 = \frac{1}{3}$ and $w_2 = \frac{2}{3}$ respectively. The respective probabilities of getting 10 dominants, given the mating, are $e_1 = 1$ and $e_2 = 1/2^{10} = 1/1024$. A substitution in (1) shows that the chance that the mating was in fact $GG \times gg$ is $w_1e_1/(w_1e_1 + w_2e_2) = 512/513$, i.e., near certainty, which is what common sense would suggest.

(b) A coin can (we assume) be either double-headed, in which case it will always come down heads, or normal, coming equally often heads and tails. We take a coin at random, toss it 10 times, and find it comes heads every time. What is the chance that it is double-headed? This is similar to the previous problem in every respect except the initial probabilities. We might perhaps know that one coin in a million is double-headed, so that now $w_1 = .000001$, $w_2 = .999999$: a calculation then would show that the chance that the coin is double-headed is only about .001. Even though the observed data are very unlikely to occur if the coin is normal, the alternative supposition that it is double-headed is still more unlikely, and so we rightly believe in the normality of the coin. In spite of the similarity of the two problems, the conclusions go in opposite directions.

(2) SIGNIFICANCE TESTS

A serious obstacle to the application of Bayes' Theorem in practice is that we often have only very vague ideas of what the initial probabilities are. Suppose, for example, we wish to know whether a given type of mating does or does not give a 1:1 segregation on the basis of the observed numbers from such matings. To decide this we would require to know the initial chance that a mating "chosen at random" gives a 1:1 segregation. Not only is this difficult to estimate, but it is also open to serious question whether the mating under study can be legitimately said to be "chosen at random." In consequence, Bayes' Theorem has fallen into disfavour and most statisticians now use "significance tests" or other alternative methods of inference.

The principle of a significance test is, in its essentials, that if we can show that the observed data are very unlikely to occur if a hypothesis H_0 is true, we can regard them as "significantly" opposed to the hypothesis, and in the absence of good reason to the contrary we will tend to disbelieve the hypothesis. Thus if the hypothesis is that a given mating should give a 1:1 segregation, the chance that it gives a family of 10 children all alike (i.e., all dominant or all recessive) is $2/2^{10} = 1/512$. If then such a family occurs, the hypothesis of a 1:1 segregation is "rejected at a significance level 1/512." This appears to be the point of view of Sir Ronald Fisher, who is responsible for many of the most important tests of significance: in his book (Fisher 1956, p. 39) he says:

"The force with which such a conclusion is supported is logically that of the simple disjunction: *Either* an exceptionally rare chance has occurred, *or* the theory ... is not true."

Significance tests have the convenient property that they summarize the data in

a unique numerical manner: two persons given the same data will find the same level of significance (or perhaps nearly the same, if slightly different methods of calculation are used). But on the other hand, the interpretation of the test will depend on the background, experience, and prejudices of the investigator: if he (rightly or wrongly) believes very strongly in the hypothesis tested (H_0), he will still continue to believe in it, even if it is contradicted by the data at a high level of significance. He will say that this apparent contradiction is due either to an exceptional chance fluctuation or to bad experimentation.

When Bayes' Theorem can be applied, it is more informative than a significance test, for it gives to each hypothesis an exact probability of being true. A significance test, on the other hand, may "reject a hypothesis at significance level P", but P here is *not* the probability that the hypothesis is true, and indeed the rejected hypothesis may still be probably true if the odds are sufficiently in its favour at the start. For example, in human genetics there are odds of the order of 22:1 in favour of two genes chosen at random being on different chromosomes; so even if a test indicates departure from independent segregation at the 5 per cent level of significance, this is not very strong evidence in favour of linkage. (Morton allows for this in his sequential tests.)

(3) DECISION PROCEDURES

An alternative theory of significance tests was given by Neyman & Pearson (1933) and this had developed into Wald's (1947, 1950) theory of decisions. A simple example is that of a manufacturer of articles who wishes to test a batch of them before sending them on to a purchaser. A certain small fraction may inevitably be expected to be defective: the manufacturer wishes to keep this fraction reasonably low without going to the expense of testing every article individually. He therefore takes a sample from the batch and tests it. If the sample is satisfactory he "accepts" the batch and sends it on, if not, he "rejects" the batch, retaining it for further tests or for destruction. Naturally, this procedure cannot be absolutely infallible; unless he takes the whole batch and tests each article individually, he can not know for certain exactly how good it is. (In some cases testing involves destroying the article and so such a complete test of a batch is in any case useless.) Among "good" batches, defined as those with only a certain small proportion of defectives (e.g. 1 per cent) there will be a few which will fail the test purely by chance, be rejected, and so involve the manufacturer in further expense, such as in testing the whole batch. Such a misleading rejection is called an "error of the first kind" and the proportion of good articles which are rejected in this way is usually denoted by α. Similarly among "bad" batches, (e.g. those with 5 per cent or more of defectives) a small proportion will pass the test by chance, and will cause trouble to the consumer. These are "errors of the second kind," and the proportion of bad batches which are misleadingly accepted is usually denoted by β. Both kinds of error are undesirable, so the general aim of the manufacturer will be to keep α and β as low as is possible without unreasonably increasing the cost of testing.

A test of this kind, which leads to a choice between two or more possible courses of action, is generally known as a "decision procedure".

(4) SEQUENTIAL TESTS

The simplest kind of decision procedure is the one considered above, in which batches have to be sorted into two categories, "good" and "bad." Wald (1947) showed that the most efficient way to proceed in such cases was not to take the whole sample at once and test it, but rather to draw items one by one until a sufficient number of good or bad articles have been drawn to make it clear which way the decision should go. Roughly speaking, this may be compared to the scoring system in tennis as contrasted with that in football. Tennis proceeds "sequentially," in that play goes on until one player has managed to get a definite lead of a certain number of points over the other (although the rules defining this lead are somewhat complicated in detail). The time needed to complete any match therefore depends on the relative strengths of the players; if they are nearly equal it may be some time before a final decision is reached, but such a decision indicates the better player with reasonable reliability. A football match is normally a "fixed sample size procedure," in that it goes on for a definite time and then ends. If the teams are very unequal, it may be clear which is the better one early in the game, whereas if they are nearly equally matched the final decision may be very much a matter of chance.

From the point of view of the manufacturer sequential tests have very much the same advantage of reaching a reliable conclusion as quickly as possible. The argument considered above suggests that this will be so if the "bad" batches are heterogeneous, for then by a sequential procedure the really bad batches will be detected very quickly, economizing in time and expense of testing, while the less bad ones will take longer, but no longer than is really necessary. A fact which is not quite so obvious, but which was demonstrated by Wald (1947), is that sequential tests are best even in the cases in which all "bad" batches are equally defective. Suppose that each good batch contains a known proportion p_0 of good articles, and $q_0 = 1 - p_0$ of defectives, whereas each bad batch contains a known proportion p_1 of good articles, q_1 of defectives. The method adopted is then as follows: for each article X we calculate a "score" $z(X) = \log (P_B/P_G)$, where P_G means the probability of getting an article like X when the batch is good, and P_B the same when the batch is bad, thus when X is good $P_G = p_0$ and $P_B = p_1$ and so

$$\left.\begin{array}{l}\text{if } X \text{ is good, } z(X) = \log (p_1/p_0) < 0;\\ \text{if } X \text{ is bad, } \quad z(X) = \log (q_1/q_0) > 0.\end{array}\right\} \tag{2}$$

As the articles are drawn from the batch, their sources are added consecutively to make a running total score Z. We also choose two "goals," a positive number a and a negative number b. If at any time the total score Z reaches a, we decide to reject the batch, whereas if it falls to b, the batch is accepted. If we define as before α to be the proportion rejected out of all good batches, and β the proportion accepted of all bad batches then nearly enough $a = -\log \alpha$ and $b = \log \beta$ when α and β are sufficiently small. Thus one can design a test with predetermined values of α and β. [More accurately $a = \log (1 - \beta) - \log \alpha$ and $b = \log \beta - \log (1 - \alpha)$].

In practice one rarely meets with quite so simple a situation in which all "bad" batches are equally bad. The usual custom is accordingly to proceed roughly as

follows. Let q_0 be the proportion of defectives produced when the process is working well, i.e. in good batches. Let q_1 be the maximum proportion of defectives that the manufacturer is prepared to tolerate, i.e. any batch worse than this ought to be rejected. Write $p_0 = 1 - q_0$, $p_1 = 1 - q_1$, as before. Then we proceed with a sequential test for comparison between proportions q_0 and q_1 of defectives: the scores are calculated from (2) and the addition of scores proceeds as before until one of the goals a or b is reached. This test accordingly has the following properties:

(i) If the proportion of defectives is q_0, the batch is good: from the preceding theory we then know that the batch has a chance α of being mistakenly rejected, and $(1 - \alpha)$ of being accepted. Now in a significance test the "significance level" is defined as the probability of rejecting the hypothesis H_0 tested, when H_0 is in fact true; so by analogy it is natural to refer to α here as the "significance level" of the sequential test, i.e. the probability of mistakenly rejecting a batch when it is in fact good.

(ii) If the proportion of defectives is in fact q_1, the chance of accepting the batch is β and of rejecting it is $(1 - \beta)$. Since β is ordinarily chosen to be a small number, the majority of such batches will be rejected.

(iii) If the proportion of defectives is greater than q_1, the batch is definitely bad. Since each defective scores positively the score will normally tend to the positive goal a even more rapidly than if the proportion was equal to q_1; so on the average the batch is rejected even more quickly, and the chance of it being mistakenly accepted is even smaller. Thus β is an upper bound to the proportion among all bad batches of those which are mistakenly accepted ("errors of the second kind").

(iv) If the proportion of defectives is between q_0 and q_1 the batch is tolerable, even if not completely good. In such a case the manufacturer will not be over much concerned whether the batch is accepted or rejected.

This procedure seems therefore to give very reasonable results. The test used will be the one best adapted to distinguish between the two proportions q_0 and q_1 of defectives: it will therefore not be absolutely the best test for the real situation in which we get varying proportions of defectives in the bad batches. However, the computations in this simplified test are so easy that it is usual to ignore this slight loss of efficiency in the interests of ease of calculation. One could devise a sequential test which would take into account the varability of the proportion of defectives, but it would be rather more cumbersome, and we will not discuss it here. The same applies to a "closed sequential test" in which the sampling proceeds sequentially up to a certain fixed size, and then stops.

SEQUENTIAL TESTS APPLIED TO LINKAGE PROBLEMS

Morton (1955) has proposed that sequential tests should be applied to the detection of linkage. This can very readily be done by a verbal adaptation of the preceding argument. Instead of drawing a sample of individuals from a batch of goods, we collect families from a population, and study a particular pair of (autosomal) loci. Instead of asking whether the batch is "good" or "bad" we ask whether the loci are respectively unlinked or linked. Instead of saying that some batches are so nearly good as to be tolerable, Morton argues that some pairs of loci are so loosely linked that it would not be practicable to try to detect the linkage in any reasonably

attainable body of data. We therefore choose a value θ_1 of the recombination fraction which represents the loosest linkage we can reasonably expect to detect, and compare this with no linkage, i.e. recombination fraction $\frac{1}{2}$. The score for any family F is defined to be

$$Z(F) = \log \text{Prob}\,(F|\theta_1) - \log \text{Prob}\,(F|\tfrac{1}{2}) \qquad (3)$$

where Prob $(F|\theta_1)$ is the probability of occurrence of the family F when the recombination fraction is θ_1, and Prob $(F|\frac{1}{2})$ is the similar probability in the absence of linkage. These scores for separate families are added successively until the total score reaches either the positive goal a, indicating the presence of linkage, or the negative goal b, indicating its absence. For details of the method of calculation of scores, and tabulations of their values in various commonly occurring types of family see Morton (1955, 1957).

THE PRINCIPLE OF SELF-SUFFICIENCY

Suppose we have taken at random a sample S of individuals from a population, and wish to consider what information this gives about the population. To take a simple example, suppose we are interested in the proportion q of defectives, and that the sample contains n individuals, of which x were found to be normal and y defective. The information given by the sample may be considered to have three aspects.

(i) The numbers x, y of normals and defectives in the sample;

(ii) The order in which the normals and defectives occur (e.g.) NDDND or DNNDD etc.);

(iii) The fact that the sampling stopped at that point, and did not continue further.

Now once we are given the numbers x, y of normals and defectives respectively, the order in which they occur is purely a matter of chance. Some arrangements may be ruled out by the stopping rule, in that if such an order of normals and defectives has occurred, the sampling would have been brought to an end at some previous stage. But all arrangements which are not ruled out in this way are equally likely to occur, and so the particular order observed gives no information about the proportion q of defectives (though it may give information about the stopping rule). As for the fact that sampling has been stopped at this particular point, it seems clear that this can contribute no additional information about q, since it is essentially a decision entirely at the discretion of the investigator. If he bases it on some stop rule which depends on the number or order of occurrence of normals and defectives, this merely uses the information (i) and (ii) over again; if he bases it on any other considerations it will be irrelevant to the issue. (Any extra-sensory perception of the correct answer would have to be considered as further sampling, whereas we assume that sampling has stopped!) This intuitive kind of argument can be extended to more complicated cases, and shows that so long as sampling is at random, the information which any sample provides about the population from which it is drawn is contained entirely in the numbers of the different types of objects occurring in the sample. The order in which they occur, and the stop rule which ends the sampling, are both completely irrelevant. We might call this the "Principles of Self-sufficiency of the Sample."

It seems a reasonable deduction from this that the statistical analysis of the results should depend only on the numbers observed, and not on the order of occurrence or on the stopping rule. It does not matter whether the experimenter has finished because of a sequential, closed sequential or fixed size stop rule, or because he has exhausted all available data, or because he is just bored, the conclusions to be drawn from his data should be the same, otherwise the statistician is being inconsistent, in drawing sometimes one conclusion and sometimes another from the same information.

These conclusions are in harmony with various investigations: for example, Anscombe (1953, 1954) in papers on sequential estimation, points out that estimators used in ordinary statistical theory for samples of fixed size will apply equally well to large samples obtained sequentially: e.g. the estimate of the proportion q of defectives will still be y/n, with standard error $\sqrt{(pq/n)}$ (where $p = 1 - q$) as usual. Barnard (1953) points out that the maximum likelihood estimates do not depend on the stop rule. Wald (1950) shows that the "best" decision procedures must be of the so-called "Bayes" type, which means that they are calculated from the likelihood function for the sample, which does not depend on the stop rule. (In technical terminology, the Bayes decisions form a complete class of admissible decision functions.)

Unfortunately, if this argument is accepted it leads to some rather startling conclusions. Very few statistical procedures in general use at present are valid independently of the stop rule. Significance tests are a particularly awkward exception; if the principle of self-sufficiency of the sample was to apply to them, it would mean that we should be able to apply in all cases the ordinary significance tests designed for a sample of a fixed size. However, it is very well known that one can often cheat (intentionally or otherwise) by continuing to gather data until a particularly favourable chance fluctuation occurs, and then stopping and applying the usual significance test to show that this fluctuation is "significant". Anscombe (1954) shows how to "prove" or "disprove" any given statistical hypothesis in this way. Of course, this does not mean that ordinary statistical procedures do not work reasonably well when used properly; but it does suggest rather strongly that they are not entirely satisfactory, and it is conceivable that in due course some new methods may be put forward which do not have these disadvantages, or lead to these paradoxical results in certain situations, and which will gain general acceptance. I shall show later, however, that very fortunately in the particular case of linkage detection the difficulties can be overcome without introducing any new or unconventional methods.

What precisely is the advantage of a sequential procedure? It cannot extract any more information from a given sample than can fixed-size sampling, for as we have seen, the information depends only on the sample and not on the stop rule. It seems therefore that the advantage is simply that of finishing sampling when we have collected sufficient data for the purpose in hand. If the object is to accept or reject a batch of goods, as in the industrial case already considered, Wald's stop rule may be the appropriate one. But it is not difficult to think of cases where other stop rules would be more appropriate. For example, the proportion q of defectives in the sample will be estimated as y/n, and the standard error of this estimate will be

nearly enough $\sqrt{(pq/n)} \simeq \sqrt{(xy/n^3)}$. If therefore we wish to estimate the proportion q with some definite accuracy, i.e. with a standard error not exceeding some given value S_0, we must continue collecting data until $\sqrt{(xy/n^3)} < S_0$, but will then presumably stop as the required accuracy has been attained.

COMMENTS ON MORTON'S SEQUENTIAL LINKAGE TEST

In the case of an industrial experiment, the object is to make a choice between two definite courses of action, namely the acceptance or rejection of a batch of goods. In translating this into a sequential linkage test, the corresponding object becomes verbally to "accept" or "reject" the hypothesis of linkage. These phrases "accepting" or "rejecting a hypothesis" are much used in the literature of significance test, but the question arises as to what they mean, if anything. If we "accept" or "reject" a hypothesis, we do not in general perform any actions as a direct consequence of the decision. (Such more remote problems as giving eugenic advice will be considered later.) Nor does "acceptance" mean unconditional and complete belief in the existence of linkage, since any sample can only make linkage seem more or less probable, never completely certain. The situation is that any body of data will leave us more or less inclined to believe in linkage between the loci under study, and the reasonable procedure would be to construct some measure of how convincing the data are. But this is not a decision problem in the technical sense of the word; it does not call for a definite choice between two or more alternative courses of action, and accordingly there does not seem any particular point in using a decision procedure, or a Wald sequential test. It is not easy to think of any situations in the real world which would correspond to "acceptance" or "rejection," in the sense that these are precisely defined decisions, and that the object of the investigation has been fulfilled when one or the other has happened. We could perhaps think of "acceptance" as being sufficiently strongly convinced of the existence of linkage, and "rejection" as sufficiently strong conviction of its absence, and so reduce the question formally to a decision problem. However, the operation of a sequential stop rule would then mean that one was only interested in reaching that degree of conviction; when it has been reached, one would stop sampling, and refuse to consider further evidence, turning instead to some other problem. This would be very odd behaviour if the data favored linkage; one would rather expect the investigator to gather fresh data with increased eagerness. But the Wald sequential procedure does not make any provision for continuing after the "stop." This would be rather like getting two tennis players to play a match, decide who had won, and then continue playing for some extra time to confirm the decision. That is not efficient, for if they have to play this extra time, it would be more reliable to decide who had won on the basis of the final results than on the less complete results at the intermediate point. (We assume that there is no variation in the strength of the players as a result of fatigue or practice effects, since in the analogous statistical tests the probabilities of "defective" or "linkage" are supposed to remain constant.)

In view of the above discussion it seems best to divide the questions to be answered in two parts.

(i) What information does the sample give us about linkage?

(ii) Is it worth while continuing to gather further data?

In a Wald sequential test these questions are bound together; at the same time as a decision is made in favor of or against linkage, it is also decided to stop sampling. Nevertheless these two questions are somewhat different in nature: the second one is a decision problem, calling for a choice between the possibilities of continuing the investigation or abandoning it. The first is not a decision problem, and is not really appropriately dealt with by a sequential test.

BAYES' THEOREM APPLIED TO LINKAGE ESTIMATION

Since our argument suggests that neither significance tests nor sequential tests are entirely appropriate for the analysis of linkage data, the question arises urgently whether there is any accepted statistical method which will do better. Fortunately it is possible to apply Bayes' Theorem. The basis of this is Morton's (1955) demonstration, on theoretical and empirical grounds, that we can give a reasonaby good approximation to the initial distribution of the recombination fraction θ. Since there are 22 pairs of autosomes in man, there is a probability of about $^{21}\!/_{22}$ that two autosomal genes chosen at random will be on different chromosomes; in the remaining case of linkage, the recombination fraction will have an approximately uniform probability of taking any value between 0 and $\frac{1}{2}$. The situation is directly comparable with that considered in equation (1), except that instead of finding the probability that a given event E happened through way W_r, we have to find the probability that a given sample S arose in a population with recombination fraction θ. There is an additional complication in that instead of a finite number of ways W_r we have a continuous range of values of θ. This last difficulty can be overcome with negligible loss of accuracy by taking only the first two decimal places in θ i.e. by proceeding as if θ could take only the exact values .00, .01, .02, ... .50. The initial probability $^{1}\!/_{22}$ of their being linkage must be divided equally between the 50 values from $\theta = .00$ to $\theta = .49$, giving each one an initial probability $^{1}\!/_{1100}$; the remaining value $\theta = .50$, corresponding to no linkage, has initial probability $^{21}\!/_{22}$. Let $e(\theta) =$ Prob $(S|\theta)$ denote the chance of the observed sample S arising when the recombination fraction is known to be θ; by (1) the choice that the recombination fraction is θ after the observation is S is, when $\theta < .50$

$$\text{Prob } (\theta \mid S) \frac{{}^{1}\!/_{1100}\, e(\theta)}{{}^{1}\!/_{1100}[e(.00) + e(.01) + \cdots + e(.49)] + {}^{3}\!/_{2}{}^{1}\!/_{2}\, e(.50)} \tag{4}$$

The value of Prob $(\theta = .5|S)$ has the same denominator, but the numerator is now ${}^{3}\!/_{2}{}^{1}\!/_{2}e(.50)$. Now we can simplify (4) in the following way. Write

$$\lambda(\theta) = e(\theta)/e(.50) = \text{Prob } (S|\theta)/\text{Prob } (S|.50) \tag{5}$$

Furthermore, let Λ denote the average value of $\lambda(\theta)$, values of θ lying between 0 and .49, i.e.

$$\Lambda = [\lambda(.00) + \lambda(.01) + \ldots. + \lambda(.49)]/50$$

Then if we multiply numerator and denominator on the right hand side of (4) by $50/e(.50)$ it becomes

$$\text{Prob } (\theta|S) = \lambda(\theta)/50\ [\Lambda + 21] \tag{6}$$

when $\theta < .50$. This probability means, of course, the chance that the recombination fraction takes the value θ *correct to two decimal places*. In the same way, we find that the probability that the loci are unlinked is

$$\text{Prob } (\theta = .50|S) = 21/[\Lambda + 21] \tag{7}$$

In order to use these formulas it only remains to find the values of $\lambda(\theta)$ and Λ. Now we can show that

$$\lambda(\theta) = \text{antilog } Z(\theta) \tag{8}$$

where $Z(\theta)$ is the total lod score. Suppose that the sample consists of families F_1, $F_2 \ldots . F_h$. By the multiplication law of probabilities, the probability of occurrence of the whole sample, given θ is

$$\text{Prob } (S|\theta) = \text{Prob } (F_1|\theta)\ \text{Prob } (F_2|\theta) \ldots \text{Prob } (F_h|\theta)$$

Therefore

$$\lambda\ (\theta) = \frac{\text{Prob } (S \mid \theta)}{\text{Prob } (S \mid \tfrac{1}{2})} = \left(\frac{\text{Prob } F_1 \mid \theta}{\text{Prob } F_1 \mid \tfrac{1}{2}}\right) \cdots \left(\frac{\text{Prob } (F_h \mid \theta}{\text{Prob } (F_h \mid \tfrac{1}{2}}\right) \tag{9}$$

But by definition, the score $z_r(\theta)$ for family F_r is

$$z_r\ (\theta) = \log \text{Prob } (F_r \mid \theta) - \log \text{Prob } (F_r \mid \tfrac{1}{2}) = \log \frac{\text{Prob } (F_r \mid \theta)}{\text{Prob } (F_r \mid \tfrac{1}{2})}$$

Hence by taking logarithms on both sides of (9) we get

$$\log \lambda(\theta) = z_1(\theta) + z_2(\theta) + \ldots + z_h(\theta) = Z(\theta)$$

where $Z(\theta)$ means the total score, being the sum of the scores for each family. This is equivalent to (8). Nor ordinarily it is tiresome to calculate the exact score $z_r(\theta)$ for every family, taking into account the whole of the available information. But as Morton points out, the calculations can be very greatly simplified with only a small loss of information by including those families in which the genotypes of the parents are completely known (except for phase), and applying a correction to their scores to allow for the omission of the other families. In this case we can find the values of $z(\theta)$, for $\theta = .05, .1, .2, .3$ and $.4$ from Morton's (1955, 1957) tables. For $\theta = .5$, $Z(\theta)$

TABLE 1. STEINBERG AND MORTON'S SCORES

Recombination fraction θ	0	.05	.1	.2	.3	.4	5
Total score $Z(\theta)$	$-\infty$	−2.631	−.274	.716	.528	.175	0
$\lambda(\theta)$ = antilog $Z(\theta)$	0	.002	.532	5.200	3.373	1.496	1

is always zero by definition. The value of $Z(0)$ is not given, but in any case in which there is at least one certain crossover (as will happen in the great majority of linkage investigations) the total score Z(0) is $-\infty$, and $\lambda(0) = 0$. For example, the total scores found by Steinberg and Morton (1956) for linkage between cystic fibrosis and the MNS blood groups are shown in Table 1. This only gives a few values of $\lambda(\theta)$. Although a completely accurate determination of $\lambda(\theta)$ for other values of θ would require an extension of Morton's tables, a great deal can be done without this. Λ is by definition the average value of $\lambda(\theta)$ for values of θ between 0 and $\frac{1}{2}$; this can be calculated approximately by a formula for numerical integration, such as Simpson's rule.

$$\Lambda \simeq \tfrac{1}{30}\,[\lambda(0) + 4\lambda(.05) + 3\lambda(.1) + 8\lambda(.2) + 4\lambda(.3) + 8\lambda(.4) + 2] \qquad (10)$$

This gives approximately $\Lambda = 2.355$; whence from (7) the probability that the genes are not linked is Prob $(\theta = \frac{1}{2}|S) = 21/(\Lambda + 21) = .899$. However, it is rather more accurate to use interpolation formulas to find intermediate values of $\lambda(\theta)$. The following formulas, obtained by fitting a polynomial to $Z(\theta)$, are convenient and of high accuracy.

$$\begin{aligned}
Z(.15) &= -.046Z(.05) + .345Z(.1) + 1.203Z(.2) - 1.148Z(.3) + 1.422Z(.4)\\
Z(.25) &= .007Z(.05) - .033Z(.1) + .355Z(.2) + 1.053Z(.3) - .762Z(.4)\\
Z(.35) &= -.002Z(.05) + .007Z(.1) - .043Z(.2) + .402Z(.3) + .949Z(.4)\\
Z(.45) &= .001Z(.1) + .008Z(.2) - .044Z(.3) + .373Z(.4)
\end{aligned}$$

Using these we get from Table 1 the following interpolated values:

Recombination fraction θ	.15	.25	.35	.45
Interpolated score $Z(\theta)$	.531	.667	.351	.048
$\lambda(\theta)$	3.396	4.645	2.244	1.117

By using these interpolated values of $\lambda(\theta)$ we can calculate Λ from Simpson's rule.

$$\begin{aligned}
\Lambda \simeq \tfrac{1}{30}\,[&\lambda(0) + 4\lambda(.05) + 2\lambda(.1) + 4\lambda(.15) + 2\lambda(.2) + 4\lambda(.25)\\
&+ 2\lambda(.3) + 4\lambda(.35) + 2\lambda(.4) + 4\lambda(.45) + 1]
\end{aligned} \qquad (11)$$

Thus, on substituting the values calculated above we find $\Lambda = 2.261$ whence a more accurate estimate of the probability of the loci being unlinked is $21/\Lambda + 21) = .889$. For $\theta < .5$) we find from (6) the values of Prob $(\theta|S)$ in Table 2. These values are shown in Fig. 1. (It should be remembered that Prob $(\theta|S)$ means here the probability of the recombination fraction θ correct to two places, i.e. Prob $(\theta = .20|S)$ is the chance that θ lies between .20 and .21). Other values could be obtained if desired by interpolation, but these are sufficient to show that the distribution has a peak around $\theta = .2$.

TABLE 2. PROBABILITY OF VARIOUS VALUES OF THE RECOMBINATION FRACTION

θ	.00	.05	.10	.15	.20	.25	.30	.35	.40	.45	.49
Prob $(\theta\|S)$	0	.00000	.00046	.00292	.00447	.00399	.00290	.00193	.00129	.00096	.00086

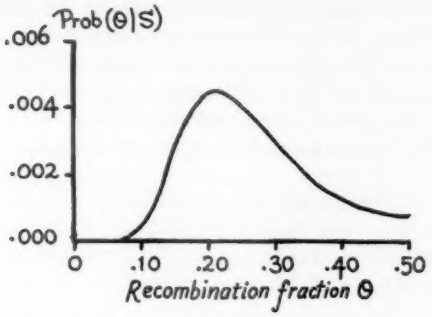

FIG. 1. Probabilities $P(\theta|S)$ of the occurrence of values of the recombination fraction θ less than .5. (There is a concentration of probability at $\theta = .5$, corresponding to the absence of linkage).

ADVANTAGES OF BAYES' THEOREM

Bayes' Theorem has the advantage that, unlike any other method of statistical inference, it gives the answer directly as a probability. Thus Steinberg and Morton's data gives a probability .9 of their being no linkage, and only .1 of linkage being present. This probability has a direct meaning, and does not need to be hedged about with qualifications, unlike a significance level. The investigator can stop sampling at any point he pleases, and for any motive, and provided he calculates the probabilities from all the data he has gathered up to that point, they will be valid. If we consider the linkage between one main character and a variety of test characters, such as blood groups, PTC taste testing, etc., it is well known that on the basis of testing by significance we can expect one test in 20 to give a significance at the 5 per cent level, or one in 100 at the 1 per cent level. The experimenter is accordingly warned not to take an occasional significance too seriously, and not to pick out the more hopeful cases and test them alone for significance. However, no such special precautions are needed in using final probabilities: each linkage test gives its own result, and can be considered entirely on its own without reference to the others. Bayes' Theorem also has the advantage of giving not only the chance of linkage being present but also the probabilities of different values of θ between 0 and .5, i.e. it combines the "detection" and "estimation" of linkage in a single calculation. With the help of Morton's (1955, 1957) tables, the amount of work required to do this is quite small—much smaller than the effort required to gather the data in the first place.

However, it must be remembered that all these useful properties depend on the availability of a valid initial distribution for θ. The one given by Morton, although very plausible, can only be approximate, and hence the probabilities calculated from it will be similarly approximate. Such a degree of vagueness seems inescapable at present; if the initial probabilities are doubtful to this extent, so also must be the final probabilities. Another source of difficulty, which will be resolved only by further

investigations, is that there may be a difference between the distributions of θ in the male and female. A certain caution is therefore necessary in interpreting the results.

Note. The interpolation formulas given above for calculating the values of $Z(.15)$, $Z(.25)$, $Z(.35)$, $Z(.45)$ were obtained on the assumption that all families were of two generations at most, with parental phases unknown. When that is not so, the formulas may be wrong and the following ones should be used instead.

$$
\begin{aligned}
Z(.15) &= -.111Z(.05) + .547Z(.1) + .729Z(.2) - .219Z(.3) + .062Z(.4)\\
Z(.25) &= .048Z(.05) - .156Z(.1) + .625Z(.2) + .562Z(.3) - .089Z(.4)\\
Z(.35) &= -.048(.05) + .141Z(.1) - .312Z(.2) + .834Z(.3) + .042Z(.4)\\
Z(.45) &= .111Z(.05) - .312Z(.1) + .583Z(.2) - .875Z(.3) + 1.250Z(.4)
\end{aligned}
$$

THE SAMPLING STOP RULE

As we have seen, when Bayes' Theorem is used, the investigator is not under any compulsion either to continue collecting data or to stop at any point merely in order to make the statistical analysis possible and valid. Nevertheless he will wish to know roughly how much data he must collect in order to have a reasonable chance of detecting linkage, and whether, after some data have been collected, it is worth while proceeding further.

Morton in his 1955 paper discussed how many families would be required to show the presence or absence of linkage according to his sequential tests. Although his methods differ in principle from ours, they will nevertheless indicate reasonably closely the order of magnitude of the number of observations required to raise the odds in favour of linkage to a reasonable value (such as 20:1); from his Fig. 4 it seems that we might reasonably need the equivalent of between 30 and 1000 double-backcross sib-pairs to do this, according to circumstances. With much less than this amount of data it scarcely seems worth beginning the calculations as the results will almost certainly be quite inconclusive.

Once the investigation has been begun, the question may arise of whether to continue it further or to abandon it. In principle this is a decision problem, and could be treated exactly by the methods of decision function theory However, there are so many complicating factors, such as the availability of data, the interest of the problem from aspects other than linkage, etc., that it seems simpler to treat the problem much more informally. The two main factors which guide the experimenter in making his decision whether to continue are (a) the chance of ultimate success (b) the probable amount of further data required to reach this success. However the chance (a) of ultimate success can be directly measured; it is the probability that the loci are linked; e.g. in Steinberg and Morton's data it is about .1. This in itself is better than the initial chance 1/22 before any data is gathered, but it is still scarcely encouraging. It is much harder to give any precise estimate of (b), because of sampling fluctuations. However, one can get a rough idea in the following way. Since the total score Z is the sum of the scores for the separate families, it will be proportional to the number of families, apart from sampling fluctuations. Thus if 3 times as much data was gathered on the linkage between cystic fibrosis and MNS, we might expect that

the Z values of Steinberg and Morton would be approximately multiplied by 3, i.e. they would become as follows:

θ =	.00	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
Z =	−00	−9.89	−.82	1.59	2.15	2.00	1.58	1.05	.52	.14	0

On this basis the probability of the existence of linkage would be found as above to be $\Lambda/(\Lambda + 21) = .61$, i.e. if the trend continues, it will then be more likely than not that linkage is present. Similarly, with scores multiplied by 5 the probability of linkage would become .96, i.e. reasonably near certainty. The situation can therefore be summarized in the statements that there is only about 1 chance in 10 of linkage being found, and that if it is present it would probably require something like 5 times the present amount of data to demonstrate it with a reasonably high probability. The investigator will be able to decide on the basis of this information whether he feels it worth continuing.

EUGENIC ADVICE

Eugenic advice can in principle be of two kinds. It may be a simple statement of the probability of a certain defect developing, or the probability that if children were born from a given mating, they would have some defect. This is an application of the calculus of probabilities, taken in conjunction with Mendelian ratios, distributions of ages of onset, maternal age effects, etc., as far as these are relevant. Alternatively it could give a straightforward recommendation not to have children in certain cases; this is of the nature of a decision problem. However, it seems best in general to confine oneself to a calculation of the probabilities, and to allow the persons involved to make the decision for themselves, in the light of all relevant circumstances.

Sometimes such a calculation depends on the value of a recombination fraction θ, when the genes responsible for the defect are linked to some markers, such as blood groups. If the value of θ is accurately known, the calculation is usually straightforward: we find a probability Prob $(D|\theta) = d(\theta)$ say, of a certain "defect" D occurring. If however the value of θ is as yet uncertain, we calculate from all available data S a probability Prob $(\theta|S) = p(\theta)$ say, that the recombination fraction takes a given value θ. (We can again suppose for convenience that θ is given to two decimal places only.) The total probability of the defect occurring in the light of all available evidence, is then

$$\text{Prob } D = \Sigma p(\theta)\, d(\theta) \tag{12}$$

Where the summation is over all values of θ. However, it would be tedious to calculate p(θ) for every θ from .00 to .50; as a rule we will find it only for θ = .00, .05, .10 ... .50 only, by intervals of .05. The value of Prob D is then given approximately by the following formula based on Simpson's rule:

$$\begin{aligned}\text{Prob } D = \tfrac{5}{3}\,[&p(.00)d(.00) + 4p(.05)d(.05) + 2(.10)d(.10) + 4p(.15)d(.15)\\ &+ 2p(.20)d(.20) + 4p(.25)d(.25) + 2p(.30)d(.30) + 4p(.35)d(.35)\\ &+ 2p(.40)d(.40) + 4p(.45)d(.45) + .6010p(.50)d(.50)]\end{aligned} \tag{13}$$

As a simple example, consider the position of a couple A ♂ × B ♀ who are contemplating having a child. A's brother E is married to B's sister F and they already have a child with cystic fibrosis, *cf cf*. If A, B, E and F are all of blood group M, while all their parents have group MN what is the probability that the child of A and B will also have cystic fibrosis?

Let θ denote the recombination fraction between MN and cystic fibrosis (with $\theta = \frac{1}{2}$ when they are not linked). E and F must both have genotypes MCf/Mcf. Since *cf* is a rare gene, A can be expected to be of genotype $Cf\ cf$ (i.e. MCf/Mcf) either if the parents of A and E are $Mcf/NCf \times MCf/NCf$ and there are two non-crossovers or if the parents are $MCf/Ncf \times MCf/NCf$, and there are two crossovers. The chance that A is $Cfcf$ is accordingly $(1 - \theta^2) + \theta^2$; the chance that B is $Cfcf$ is similarly $(1 - \theta)^2 + \theta^2$. The child will have cystic fibrosis only if both parents are $Cfcf$, and then only with probability $\frac{1}{4}$; hence the chance of a defective child D is $d(\theta) = \frac{1}{4}[(1 - \theta)^2 + \theta^2]^2$. By working out the values of $d(\theta)$ for $\theta = .00, .05, \ldots .50$, and using the values of $p(\theta) = \text{Prob}\ (\theta|S)$ already obtained from Steinberg and Morton's data (Table 2, together with $p(.50) = .889$) we find from (13) that Prob $D = .0655$. This compares with the probability $1/16 = .0625$ which would be found without taking the possible linkage into consideration. It seems that in this case, where it is doubtful whether linkage exists, the probability is scarcely altered by taking it into account. If we knew that there was close linkage, Prob D might rise as high as .25.

SUMMARY

By using Bayes' Theorem together with the initial distribution for recombination fractions established by Morton it is possible to present the results of linkage tests in probability form. This combines detection and estimation of linkage, and can be used directly for further calculation. By the use of Morton's tables of lod scores the calculations can be done very quickly, and the process has a number of advantages over sequential tests.

REFERENCES

ANSCOMBE, F. J. 1953. Sequential estimation. *J. R. Statist. Soc., Lond.* B. 15: 1–21.

ANSCOMBE, F. J. 1954. Fixed sample size analysis of sequential observations. *Biometrics* 10: 89–100.

BARNARD, G. A. 1953. (Discussion on Anscombe's 1953 paper). *J. R. Statist. Soc., Lond.* B. 15: 22–23.

BAYES, T. 1763. An essay towards solving a problem in the doctrine of chances. *Philos. Tr. R. Soc., Lond.* 53: 370–418.

FISHER, R. A. 1956. *Statistical methods and scientific inference.* Edinburgh: Oliver and Boyd.

KENDALL, M. G. 1946. *The advanced theory of statistics. vol. 2.* London: Griffin.

MORTON, N. E. 1955. Sequential tests for the detection of linkage. *Am. J. Human. Genet.* 7: 277–318.

MORTON, N. E. 1957. Further scoring types in sequential linkage tests with a critical review of autosomal and partial sex-linkage in man. *Am. J. Human. Genet.* 9: 55–75.

NEYMAN, J., AND E. S. PEARSON. 1953. On the problem of the most efficient test of statistical hypotheses. *Philos. Tr. R. Soc. Lond.* 231: 289–337.

STEINBERG, A. G., AND N. E. MORTON. 1956. Sequential test for linkage between cystic fibrosis of the pancreas and the *MNS* locus. *Am. J. Human. Genet.* 8: 177–189.

WALD, A. 1947. *Sequential analysis.* New York: Wiley.

WALD, A. 1950. *Statistical decision functions.* New York: Wiley.

Are Ovarian Pregnancies Parthenogenetic?

DAVID J. B. ASHLEY[1]

Dillon Research Fellow, Armed Forces Institute of Pathology, Washington 25, D. C.

NIDATION of the fertilized ovum in sites other than the progestational endometrium is an uncommon event and implantation in the substance of the ovary is still more rare. Isbell and Bacon (1947) calculated the frequency of ovarian pregnancy as about 1:25,000 but the total number of acceptable cases in the literature is probably not many more than 100 (Fullerton and Salmond 1958). The rarity of this condition is attributed to the necessity for the ovum to complete its maturation after it has been shed from the Graafian follicle and before it is fertilized by the spermatozoon in the Fallopian tube. While in the follicle and just after it has been released, the ovum is still a diploid cell, the reduction division necessary for it to play its part in the formation of the zygote not yet completed. The older theory of the etiology of this condition was that there was delay in release of the ovum either because of some anomaly of the follicle or because one follicle ruptured into another, allowing the ovum to mature before the second follicle, in turn, burst. The current concept is that implantation of the fertilized ovum in anomalous sites takes place in cases with some pathological condition of the Fallopian tube, causing interference with the passage of the zygote into the uterus; in most instances the anomalous implantation is in the tubal mucosa, but rarely it takes place on the surface of the ovary (Novak and Novak 1958). This concept is strengthened by the occasional observation of a decidual reaction on the surface of the ovary, even without the occurrence of pregnancy (Ober, Grady and Schoenbucher 1957).

Recently an intriguing suggestion has been made. Shettles (1957) observed cleavage of ova in intact follicles in three of a series of 400 human ovaries which he studied. In these cases cleavage could not have followed fertilization as the walls of the follicles were intact, and he concluded that this was an example of parthenogenesis. By induction from his observations he postulated that ovarian pregnancy in the human might be parthenogenetic in origin.

It was decided to investigate this hypothesis by a study of the nuclear sex of the fetal parts in a series of ovarian pregnancies because in species showing female homogamety all diploid parthenogones should be chromosomal females. Twelve cases from the files of the Armed Forces Institute of Pathology were used for this study. All conformed to the criteria accepted for the diagnosis of primary ovarian pregnancy (Novak and Novak 1958). The tubes, including the fimbrial ends, were intact and

Received January 15, 1959.

[1] Present address, David Lewis Northern Hospital, Liverpool 3, England.

Please address requests for reprints to Editorial Office, Armed Forces Institute of Pathology, Washington 25, D. C.

separate from the ovary; the gestation site was in the normal position for an ovary; and ovarian tissue was demonstrable in immediate relationship to the wall of the sac.

Sections stained by hematoxylin and eosin were available in all cases and in most further sections were cut and stained by the Feulgen technique. The nuclear sex of the fetal parts, usually chorionic villi, was determined using the criteria of Moore, Graham and Barr (1953). The epithelial cells covering the surfaces of the villi were found to be clearly defined and well suited to this purpose.

Seven of the 12 cases studied were of male nuclear sex and the other five were of female nuclear sex. In the male cases the characteristic sex chromatin body was not identified in the vast majority of cells. In the five females a sex chromatin body could readily be found in 40 to 50 per cent of the cells.

DISCUSSION

In a species which shows female homogamety parthenogenesis, if it occurs, must result in individuals carrying only X chromosomes because the female has only this type of sex chromosome. The appearance of the "sex chromatin" in the somatic cells is an indication of the presence in the nucleus of a pair of X chromosomes and should be found in all parthenogones in man. In the present series "sexing" of the fetal parts showed that of the 12 ovarian pregnancies five showed female sex chromatin and seven did not; this finding suggests that fertilization preceded the establishment of ovarian pregnancy in these cases.

Thelytoky, which is the more common mode of parthenogenetic reproduction, can be divided into two cytological types: apomictic in which there is no meiotic division in the maturation of the ovum and automictic in which the ovum undergoes reduction division and the parthenogone arises from a haploid cell. In the former case the diploid number of chromosomes is present throughout reproduction; in the latter the chromosome number is usually restored to diploidy either by endomitosis or after autofertilisation with some other cell of the oogenic line such as the polar body (Fig. 1) (White 1954).

Arrhenotokous parthenogenesis is only seen in a small number of groups of animals. In such species the male arises from an unfertilized ovum and is, ab initio, haploid although in some instances the somatic tissues attain the diploid state by endomitosis (Fig. 2).

All the forms of natural parthenogenesis are seen in invertebrates, but this mode of reproduction is not normal for any vertebrate species. Spontaneous parthenogenetic cleavage of the mammalian ovum has been reported by Shettles (1957) in the case of the human and by Strassman (1949) in cats; in these instances the cell division has not reached a very advanced stage and the largest cell mass which has been recorded only contained about 60 cells. Loeb (1932) postulated a natural parthenogenetic origin of ovarian pregnancies which he found in guinea pigs, but offered no cytological proofs. Olsen and Marsden (1954) have reported a further stage of parthenogenesis in birds. They studied turkeys and found, after incubation of eggs laid by birds which had been kept separate from the males from the age of six weeks, that approximately one in five showed evidence of cleavage and embryonic development although the onset of development was delayed when compared with that of normally

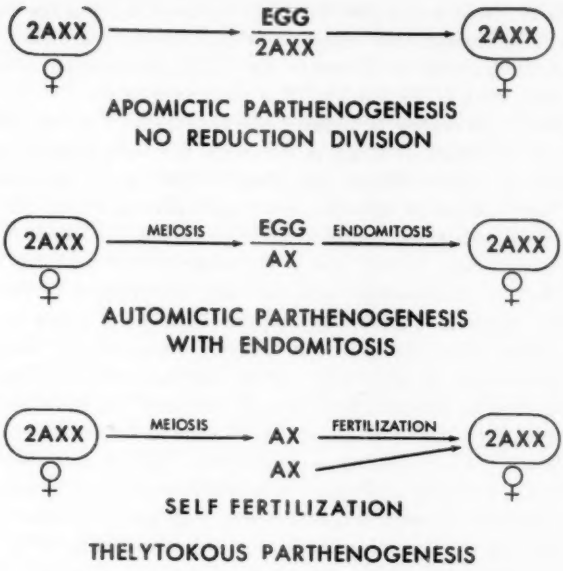

FIG. 1. Thelytokous parthenogenesis.

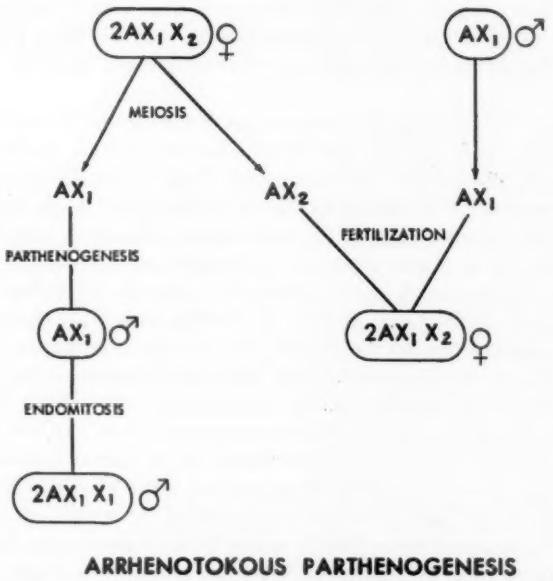

FIG. 2. Arrhenotokous parthenogenesis.

fertile eggs. In 27 instances the embryos had developed to the stage usually reached at 9 days. In these embryos the chromosome complement was shown to be diploid and in four it was possible to determine the sex by examination of the gross and microscopic appearance of the fetuses. All of these were males.

Parthenogenesis can readily be induced experimentally. Peacock (1952) cites 371 methods and, in mammals, cleavage of the ovum has been induced by cooling the Fallopian tubes in rabbits (Pincus and Shapiro 1940) and a few parthenogenetic females have been carried to full term and reared (Pincus 1951). No male parthenogones have been seen in mammalian species in which the female is homozygous for the sex chromosomes. In the case of the turkey, as in all birds, the female is heterozygous for sex chromosomes and the male homozygous. Parthenogenesis in birds, therefore, might be expected to result in embryos of either sex, females by apomictic or automictic mechanisms and males by the automictic method alone.

Objective evidence for or against the theory that human ovarian pregnancies are parthenogenetic may be obtained by a study of the sex of the embryos found in such pregnancies. Sex determination in the human is known to depend on sex chromosomes. The female has a set of autosomes and, in addition, a pair of homologous X chromosomes; the male has the same set of autosomes and an unequal pair of sex chromosomes X and Y. The larger of the male sex chromsomes is identical with the X chromosome of the female and the smaller is a special chromosome found only in the male. Fusion of the sperm and the ovum, each with a half set of chromosomes, leads to the new individual, the zygote, having a full set characteristic of the species. As half the spermatozoa carry the X chromosome and the other half the Y, the chance of any individual ovum being fertilized by a sperm carrying X or Y are equal, provided that the two types of sperm are of equal motility and viability. The primary sex ratio should therefore be approximately 1:1 with equal numbers of males and females.

In the event of parthenogenesis in man none of the embryos could possibly have a Y chromosome as the female does not have such a chromosome. Several chromosome constitutions are possible for a parthenogone. If the ovum divides after reduction division has occurred, the embryo will be of haploid constitution (automictic) and will contain half the usual number of chromosomes with one X only. Such cases do occur in nature, but as a general rule the normal chromosome number for the species is restored by endomitosis. A second alternative is that the unfertilized ovum divides by mitosis (apomictic) and that all the cells derived from it are of exactly the same chromosomal composition as the mother cell. In this case all the daughter cells must be female as is the mother cell. In the case of parthenogenesis due to fertilization of the ovum by another maternal cell the zygote would have either a diploid chromosome number, if both the cells were in the post-meiotic state, or a tetraploid chromosome number if both cells were not yet haploid. In the former instance there would be two X chromosomes in the cells of the zygote and in the latter four X chromosomes.

The method of sex diagnosis by the examination of the resting cell nuclei has been well established in man (Lennox 1956) and it is believed that the sex chromatin body seen is the resting form of the pair of X chromosomes of the female, although the absence of the sex chromatin body does not imply XY. Assuming this to be true, all

parthenogones produced by a human female would be of female nuclear sex because, whatever their particular mode of origin, in each cell there are at least two X chromosomes. The finding of the female pattern of nuclear chromatin in only five of twelve ovarian fetuses and the male chromatin pattern in the other seven is evidence that ovarian pregnancy in man is not parthenogenetic and that fertilization of an ovum by a spermatozoon with subsequent implantation must precede the development of ovarian pregnancy as it does uterine and tubal pregnancy.

Two alternative hypotheses for the absence of the female sex chromatin mass from the cells of parthenogenetic ovarian embryos must be considered. The first, that the cells of such embryos are of haploid constitution, is considered unlikely. The cells of the embryos, both of male and of female nuclear sex, were indistinguishable apart from the sex chromatin mass. Visual examination of the Feulgen stained preparations showed that the amount of D.N.A. in each type of cell was approximately the same. Were the cells haploid they would be expected to contain less nuclear material than diploid cells, the absence of such a cytological distinction is regarded as presumptive evidence against the hypothesis of haploidy. An additional point is that nothing is known of the way in which the solitary sex chromosome would be made manifest in a haploid cell. No such cells, except for spermatids and spermatogonia, are found in man and it is not certain that the nuclear sex of such cells would be male. The second alternative is that chromosome deletion has occurred and that these cells are of the sex chromosome constitution XO. This cytological anomaly has been described in a case of Gonadal Dysgenesis, Turner's syndrome, (Ford, Jones, Polani, deAlmeida and Briggs 1959) but has not yet been confirmed by other workers. It is possible, although in my opinion unlikely, that deletion of this type could occur in parthenogenetic ovarian pregnancy. Proof must wait upon further cytological examination of material from ovarian pregnancies by the most refined technics.

SUMMARY

The theory that ovarian pregnancy in the human might be parthenogenetic in origin was investigated by determining the nuclear sex of the fetal parts of twelve such pregnancies. Seven were of male nuclear sex and five of female nuclear sex. From this evidence it is concluded that ovarian pregnancy is not parthenogenetic and that fertilization must precede ovarian implantation of a zygote. The two alternative suggestions, that male nuclear sex in these cells is due to haploidy or to chromosome deletion, are considered. Haploidy is thought to be unlikely. Chromosome deletion is regarded as possible but could only be proved by further investigation.

REFERENCES

FORD, C. E., JONES, K. W., POLANI, P. E., DEALMEIDA, J. C. AND BRIGGS, J. H. 1959. A sex chromosome anomaly in a case of Gonadal Dysgenesis (Turner's syndrome). *Lancet* i: 711–713.

FULLERTON, J. M., AND SALMOND, M. 1958. Case of primary ovarian pregnancy. *Brit. M. J.* 1: 932–933.

ISBELL, N. P., AND BACON, W. B. 1947. Primary ovarian pregnancy. *Am. J. Obst.* 54: 329–335.

LENNOX, B. 1956. Nuclear sexing: A review incorporating some personal observations. *Scot. M. J.* 1: 97–121.

LOEB, L. 1932. The parthenogenetic development of eggs in the ovary of the guinea pig. *Anat. Rec.* 51: 373–408.

MOORE, K. L., GRAHAM, M. A., AND BARR, M. L. 1953. The detection of chromosomal sex in hermaphrodites from a skin biopsy. *Surg. Gynec. Obst.* 96: 641–648.

NOVAK, E., AND NOVAK, E. 1958. *Gynecological and Obstetric Pathology*. 4th edition, Philadelphia: W. B. Saunders, p. 482.

OBER, W. B., GRADY, H. C., AND SCHOENBUCHER, A. K. 1957. Ectopic ovarian decidua without pregnancy. *Am. J. Path.* 33: 199–217.

OLSEN, M. W., AND MARSDEN, S. J. 1954. Natural parthenogenesis in turkey eggs. *Science* 120: 545–546.

PEACOCK, A. D. 1952. Some problems of parthenogenesis. *Advance Sc.* 9: 134–148.

PINCUS, G. 1951. Fertilization in mammals. *Sc. Am.* 184 (3): 44–47.

PINCUS, G., AND SHAPIRO, H. 1940. The comparative behavior of mammalian eggs *in vivo* and *in vitro*. VII. Further studies on the activation of rabbit eggs. *Proc. Am. Philos. Soc.* 83: 631–644.

SHETTLES, L. B. 1957. Parthenogenetic cleavage of the human ovum. *Bull. Sloane Hosp.* 3: 59–61.

STRASSMAN, E. O. 1949. Parthenogenetic development of the ovum as observed by vital staining. *Am. J. Obst.* 58: 237–244.

WHITE, M. J. D. 1954. *Animal cytology and evolution*. 2nd edition, Cambridge University Press, p. 326 ff.

Familial Hypophosphatemia with Vitamin D-Resistant Rickets. II:

Three Additional Kindreds of the Sex-Linked Dominant Type with a Genetic Analysis of Four Such Families[1]

JOHN B. GRAHAM,[2] VERNON W. McFALLS,[3] AND ROBERT W. WINTERS[4]

From the Departments of Pathology, Pediatrics, and Medicine, The University of North Carolina School of Medicine, Chapel Hill

INTRODUCTION

THE WIDESPREAD PROPHYLACTIC USE of vitamin D in infant nutrition has virtually eliminated rickets due to deficiency of vitamin D from the United States. One consequence of this achievement has been the recognition of rickets which develops in spite of vitamin D prophylaxis. Several types of rickets of this sort have been delineated, and each type appears to result from a congenital or acquired metabolic defect (2–4, 7, 8, 10–12).

The most common form of "endogenous" rickets is an entity usually referred to as "vitamin D resistant (or refractory) rickets" (1). The clinical picture of this disorder is quite similar to that of deficiency rickets, except that large, even massive, doses of vitamin D are required for healing of the bone lesions. The only clearly defined metabolic defect is hypophosphatemia which has been attributed in large part to deficient reabsorption of phosphate from the glomerular filtrate by the renal tubular cells. It is not clear whether the latter abnormality results from an intrinsic defect in the kidney itself, or whether it is the consequence of excessive parathyroid stimulation secondary to deficient absorption of calcium from the gastro-intestinal tract. The clinical, pathological and biochemical features of this disease and the accumulated literature have been reviewed by us recently (14).

Many authors have recognized the frequent familial occurrence of "vitamin D resistant rickets". Our review of the previously published genetic data suggested that the generally accepted idea of autosomal dominant transmission might be erroneous. In the first kindred reported by our group ("E") clear evidence of a sex-linked dominant mode of inheritance was found. This result was obtained by using serum inorganic phosphorus concentration rather than skeletal abnormality to indicate presence of the abnormal gene and by analyzing the progeny of affected persons

Received April 13, 1959.

[1] Supported (in part) by research grants (H-1333 and H-3140) from the Public Health Service and a grant from the United Medical Fund of North Carolina.

[2] Professor of Pathology, University of North Carolina.

[3] Intern in Pediatrics, University of Arkansas Medical Center; formerly Student Fellow in Pathology, University of North Carolina.

[4] Assistant Professor of Physiology, University of Pennsylvania; formerly Fellow of the National Foundation for Infantile Paralysis, University of North Carolina.

separately by sex of affected parent. Penetrance was found to be high (possibly complete) for hypophosphatemia, but considerably reduced for overt bone lesions (13, 14).

Many hypophosphatemic persons in the "E" kindred lacked clinical, radiographic or historical evidence of skeletal disease. This dissociation of hypophosphatemia and bone lesions usually occurred in females. Hypophosphatemic males, on the other hand, almost invariably had skeletal involvement ranging from moderate to marked rickets, post-rachitic deformities, or osteomalacia.

We have studied three similar kindreds of the sex-linked dominant variety ("B", "H", and "L") since our original kindred was reported. The results of these later studies are presented in this paper. An additional kindred in which the disease appears to have been due to a different mechanism will be reported separately (15).

METHODS

The clinical and radiographic diagnoses of active rickets or post-rachitic deformities were made using criteria outlined previously (14). Briefly, there are three major ones: (a) typical onset and progression of characteristic skeletal deformities,

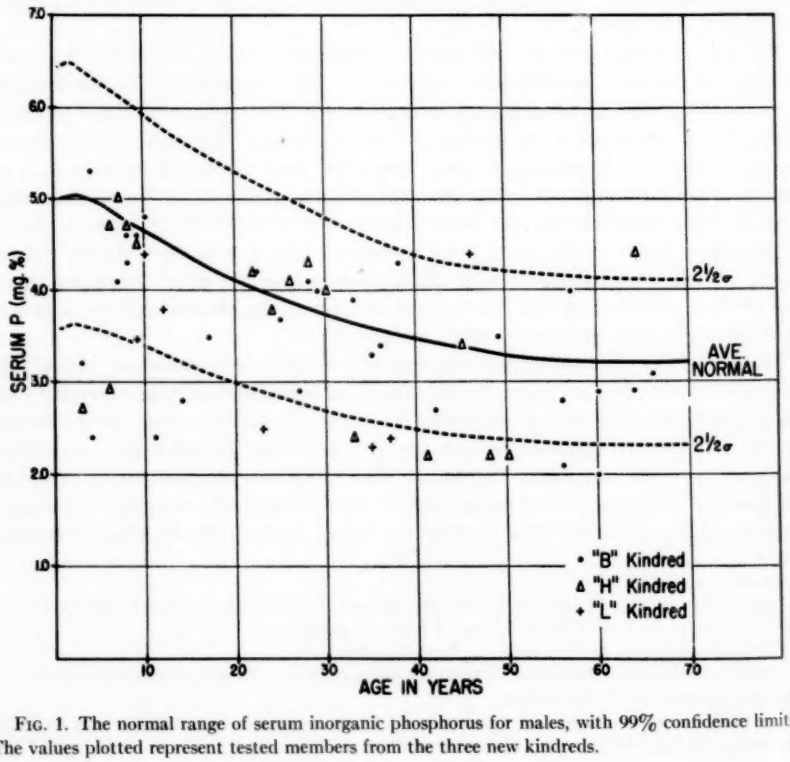

FIG. 1. The normal range of serum inorganic phosphorus for males, with 99% confidence limits. The values plotted represent tested members from the three new kindreds.

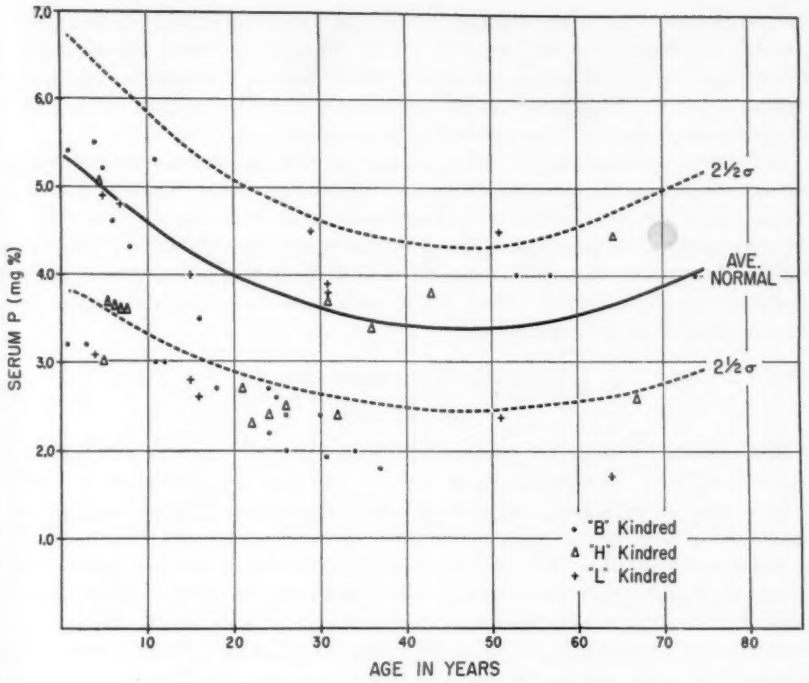

FIG. 2. The normal range of serum inorganic phosphorus for females, with 99% confidence limits The values plotted represent tested members of the three new kindreds.

despite amounts of vitamin D adequate to prevent or cure deficiency rickets, (b) typical physical signs of active rickets or of post-rachitic deformities—most often *genu valgum* or *genu varum*—usually accompanied by short stature, (c) typical radiographic signs of rickets in children or post-rachitic deformities in adults.

Nearly all of the chemical data reported in the present paper were performed on blood samples obtained before breakfast. In the few instances when this was not possible, blood samples were taken at least 4 hours after the last meal. The technique for handling blood samples and our chemical methods have been described elsewhere (5, 14). The individual determinations recorded in the tables represent averages of all values on a given patient for a particular determination. The number of such determinations ranges from one to five.

The discriminant functions used in the diagnosis of familial hypophosphatemia have been described elsewhere (5, 14). They were devised by fitting regression curves of serum inorganic phosphorus on age separately by sexes to a large normal population. An individual is considered hypophosphatemic for our purposes if his serum inorganic phosphorus falls below the lower 99% confidence limit of the normal function for his sex. Mathematical analysis has shown that this lower confidence limit has *average* operating characteristics of no more than about 0.5% false positives and 6%

false negatives, even if the frequency of familial hypophosphatemia in the general population is assumed to be as great as 5% (5). If, as is more likely, the incidence is 1% or less, the discriminant is considerably more efficient. It should be emphasized, however, that the discriminant is not uniformly efficient, being more efficient with adults than children and least efficient with prepubertal girls (5).

In Figures 1 (males) and 2 (females) we have plotted the normal regression curves for serum inorganic phosphorus with their 99% confidence limits. Against these backgrounds have been plotted the serum phosphorus values of all tested members of the three kindreds to be described below. Those persons whose values were below the lower broken curve (lower 99% confidence limit of the normal population) have been scored as hypophosphatemic. Those whose levels were above this lower curve have been scored as normophosphatemic.

EXPERIMENTAL OBSERVATIONS

"B" Kindred

History of the Proband: The proband (IV-45 in Figure 3) was a 29 year old white female brought to our attention by an associate who knew of our interest in familial rickets. She was the product of a normal full-term pregnancy and delivery and was considered entirely normal until a *genu varum* deformity developed when she began to stand and walk alone. This deformity persisted despite a total of approximately 60 ml. of natural cod-liver oil ingested over a period of several years. Radiographic signs of active rickets were present at the knees at 5 years of age. Corrective osteotomies were attempted at this time with only partial success, but the deformities stabilized following surgery. She had had two pregnancies, both deliveries requiring Caesarian section because of a rachitic pelvic deformity.

Physical examination (Table A) revealed a well-developed woman, 57 inches tall. The only significant deformity was a moderate *genu varum*. Skeletal x-rays showed a coarse trabecular pattern in the long bones, particularly the legs. Pseudofractures and generalized osteomalacia were absent, but there was moderate contraction of the pelvic inlet.

Urinalysis was negative. Chemical examination of her serum revealed only one abnormality, hypophosphatemia (see Table B).

Description of the Kindred: Figure 3 is a pedigree chart summarizing the data on 208 members of the "B" kindred. The great majority of these persons are farmers and live in Pitt County on the Coastal Plain of North Carolina. One hundred eighty-nine persons were examined by one of us, and the serum inorganic phosphorus was measured on 53.

The clinical information on 20 members of the "B" kindred found to be hypophosphatemic is tabulated in Table A and the chemical information in Table B of the Appendix. Information on 33 normophosphatemic members of the kindred is contained in Table C. A total of 11 individuals, 6 male and 5 female, had active rickets or post-rachitic deformities on clinical and radiographic grounds. In addition, histories highly suggestive of typical deformities were obtained on III-2, III-3 and III-11 who are dead. Equivocal or sometimes conflicting histories of deformities were obtained on II-1 and II-2.

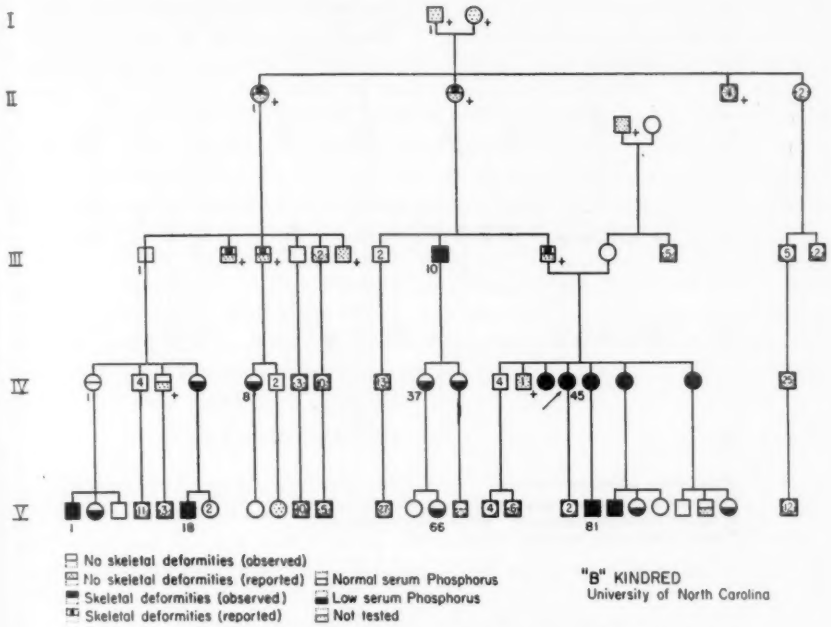

FIG. 3. The "B" kindred with sex-linked dominant hypophosphatemia.

All of the 11 persons with bone disease still living had serum inorganic phosphorus levels which were abnormally low, judged by our discriminant. In addition, 9 other persons (all women) closely related to those with bone lesions had unequivocal hypophosphatemia without clinical, radiographic, or historical evidence suggesting bone disease.

From the point of view of bone disease alone, penetrance is considerably reduced in this kindred. However, when the pedigree chart is scrutinized for hypophosphatemia, every hypophosphatemic person in generations IV and V is seen to have had either a hypophosphatemic parent or one with bone disease.

It is noteworthy that in this kindred, as in the one previously reported (13, 14), all 6 of the hypophosphatemic males had clinically recognizable rachitic bone disease, whereas only 5 of 14 hypophosphatemic females had recognizable skeletal abnormality (see Table A). It is of interest that all 5 of the hypophosphatemic women *with* bone disease were in the sibship of the proband, a female, and that their affected parent was the father. All had had moderately severe rickets as children (verified in 4 instances by examination of old roentgenograms) and were observed to have post-rachitic deformities as adults.

"H" Kindred

History of the Proband: The proband (IV-2, Figure 4) was a four year old white boy referred because of *genu varum* refractory to treatment with vitamin D. He

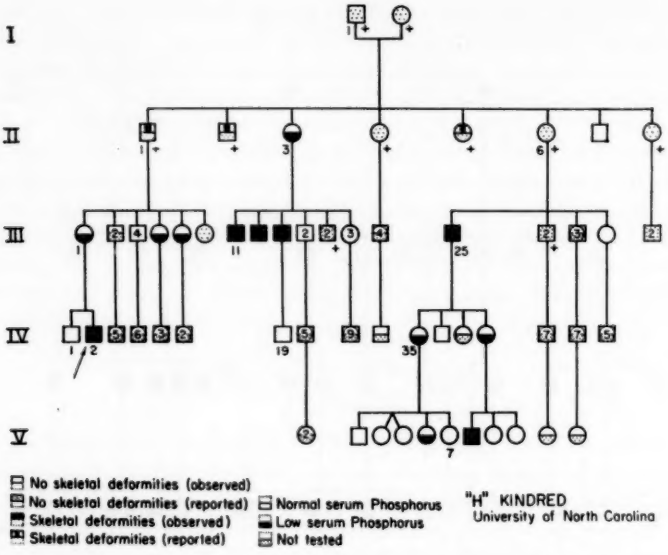

FIG. 4. The "H" kindred with sex-linked dominant hypophosphatemia.

was born after a normal, full term pregnancy and delivery, and his growth and development during early life had been entirely normal. His legs were noted to be bowed when he began walking at the age of 14 months. He had been given 5000 I.U. of vitamin D per day and liberal exposure to sunlight after the diagnosis of rickets had been made. At the age of three years his parents noted frontal enlargement of the head and a deformity of the sternum. The *genu varum* deformity progressed for 3 years despite heavy vitamin D therapy.

On physical examination at age 4 years, he weighed 37 pounds (50th percentile) and was 38¼ inches tall (<3rd percentile). Prominent frontal, parietal and occipital bossing of the head were present, but the teeth were normal. A prominent "pigeon breast" deformity of the chest was present along with Harrison's grooves and slight enlargement of the costochondral junctions. There was marked *genu varum* deformity and slight anterior bowing of both tibiae. The epiphyses were enlarged at the knees.

Radiographic studies revealed moderately severe rickets, particularly at the junctions of the lower femoral metaphyses and epiphyses, but only minimal evidence of rickets was observed at the wrists and ankles. The femoral shafts showed coarse trabeculation.

The hemogram and urinalysis were normal, and the Sulkowitch test was negative. Clinical data on the patient are shown in Table D while chemical determinations on his serum are shown in Table E. The only chemical abnormalities were the hypophosphatemia and the slightly elevated serum alkaline phosphatase.

Description of the Kindred: This patient and his kin are shown in Figure 4. Most of

the members of the family reside in Guilford and Alamance Counties, in the central Piedmont region of North Carolina, and are employed either in farming or industry. The pedigree chart shows 112 members of the kindred, 97 of whom were examined. On clinical grounds alone, 6 persons (all male) were diagnosed as having bone disease. Highly suggestive histories were obtained on three persons now dead (II-1, II-2 and II-5). The serum inorganic phosphorus was measured one or more times on 33 members of this kindred, 13 of whom proved to be hypophosphatemic. Clinical data on the 13 hypophosphatemic relatives are tabulated in Table D and blood chemical data in Table E. Data on 20 normophosphatemic relatives can be found in Table F.

As in the "B" kindred, all persons with clinical bone disease were hypophosphatemic. In the "H" kindred, however, *all* persons with clinical bone disease were males. There were 7 individuals, all female, found to be hypophosphatemic without skeletal deformities. As in the "B" family, penetrance for bone disease was incomplete, but every hypophosphatemic person was found to have a hypophosphatemic parent when this information was available.

"L" Kindred

History of the Proband: The proband (IV-1, Figure 5), a 37 year old white male, was referred because of a family history of "bowed legs". As far as could be determined, he had had a normal infancy. However, during early childhood, the *genu varum* deformity had been noted. This had progressed until adolescence when slight improvement had occurred. He had not received deliberate anti-rachitic prophylaxis or therapy, but had always been generously exposed to sunlight.

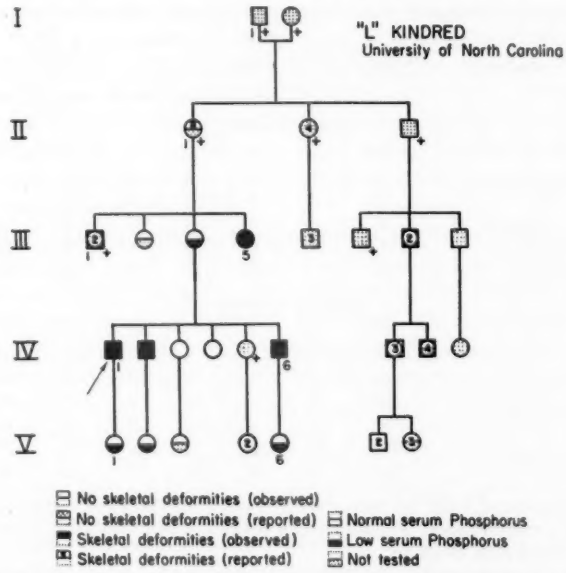

FIG. 5. The "L" kindred with sex-linked dominant hypophosphatemia.

Physical examination (Table G) revealed a short but well developed man, only 63 inches tall. There were no deformities of the head, chest or upper extremities. The teeth were markedly carious, and many were missing. The only discernible abnormality of the skeleton was a moderate degree of *genu varum*.

Urinalysis was normal. Hypophosphatemia was the only abnormality noted in his serum (Table H).

Description of the Kindred: Some information was obtained on 45 members of the kindred (Figure 5). Nearly all of these persons live in Edgecombe and Wilson Counties on the Coastal Plain of North Carolina and are farmers. Physical examination of 28 individuals in this family uncovered 3 additional cases of bone disease, 2 males and 1 female. Serum chemical studies of 19 individuals closely related to the proband showed 7 additional instances of hypophosphatemia. The 8 hypophosphatemics consisted of the 4 chemically abnormal persons, plus 4 females without clinical evidence of deformities. The clinical and chemical information on the members of the kindred are tabulated in Tables G, H and I. Once again it can be seen from the symbols on the pedigree chart that each hypophosphatemic child had a hypophosphatemic parent, but that there was "skipping" with respect to bone disease.

ANALYSIS OF THE FAMILY DATA

These three new kindreds add 15 affected males and 26 affected females to those previously reported. The amount of new information added by each of the new kindreds is not as great as by the original "E" kindred, because the new kindreds were not studied as exhaustively. It should be emphasized that the same excess of normal

TABLE 1. CLASSIFIED CHILDREN OF HYPOPHOSPHATEMIC PARENTS, SIBSHIPS COMPLETELY CLASSIFIED

Line No.	Kindred	Number of Sibships	Children						Total Children
			Males			Females			
			Affected	Normal	Total	Affected	Normal	Total	
			Affected Fathers						
1	"B"	0	0	0	0	2	0	2	2
2	"H"	1	0	1	1	0	0	0	1
3	"L"	3	0	0	0	3	0	3	3
4	Subtotal	4	0	1	1	5	0	5	6
5	"E"	5	0	9*	9	10*	0	10	19
6	Total	9	0	10	10	15	0	15	25
			Affected Mothers						
7	"B"	7	4	2	6	3	6	9	15
8	"H"	3	2	2	4	1	5	6	10
9	"L"	0	0	0	0	0	0	0	0
10	Subtotal	10	6	4	10	4	11	15	25
11	"E"	9	8	6	14	5†	9†	14	28
12	Total	19	14	10	24	9	20	29	53
Grand Totals		28	14	20	34	24	20	44	78

* The baby daughter of V-153 has been classified completing this sibship which was incomplete when "E" kindred was published. The sibship thus consists of three normal sons and two affected daughters in the birth order NNNAA.

† Number VI-69, "E" kindred originally classified as normal has been reclassified as hypophosphatemic.

TABLE 2. CLASSIFIED CHILDREN OF HYPOPHOSPHATEMIC PARENTS, *both* COMPLETELY CLASSIFIED AND INCOMPLETELY CLASSIFIED SIBSHIPS INCLUDED

Line No.	Kindred	Number of Sibships	Children						Total Children
			Males			Females			
			Affected	Normal	Total	Affected	Normal	Total	
			Affected Fathers						
1	"B"	2	0	4	4	4	0	4	8
2	"H"	2	0	2	2	2	0	2	4
3	"L"	3	0	0	0	3	0	3	3
4	Subtotal	7	0	6	6	9	0	9	15
5	"E"	7	0	10	10	12*	0	12	22
6	Total	14	0	16	16	21	0	21	37
			Affected Mothers						
7	"B"	8	4	3	7	4	6	10	17
8	"H"	4	5	4	9	1	8	9	18
9	"L"	1	3	0	3	0	2	2	5
10	Subtotal	13	12	7	19	5	16	21	40
11	"E"	9	8	6	14	5†	9†	14	28
12	Total	22	20	13	33	10	25	35	68
Grand Totals		36	20	29	49	31	25	56	105

* † See footnotes at bottom of Table 1.

males and affected females as in the "E" kindred (14) is also present in the new kindreds. In the analysis which follows, the new kindreds have been combined with the "E" kindred to give maximum numbers for the various tests.

Mode of Inheritance: The classified children of hypophosphatemic parents are tabulated in Tables 1 and 2. Table 1 includes only sibships in which all children were classified, while Table 2 includes classified children from incompletely studied sibships as well. It can be seen that the hypophosphatemic and normal sons and daughters of affected parents occur in an approximate 1:1:1:1 ratio in both tables. The ratio of affected males: normal males: affected females: normal females is 14:20:24:20 in completely studied sibships (Table 1) and 20:29:31:25 if the incompletely studied sibships are also included (Table 2).

The evidence that this undoubtedly dominant trait is not autosomal is seen in Line 6 of either Table 1 or Table 2. Affected fathers in all 4 kindreds produced only normal sons and affected daughters. The probability that this might have occurred by chance with an autosomal dominant is very slight. The χ^2 for Table 1 equals 25, and for Table 2 equals 37. (The exact probabilities are $(\frac{1}{2})^{25}$ and $(\frac{1}{2})^{37}$ respectively.) Such a distribution of children is, of course, expected with a sex-linked dominant. It should be emphasized that no instance of male-male transmission of hypophosphatemia was observed in any of the kindreds.

Penetrance: It was noted in the "E" kindred that while hypophosphatemia appeared to be fully penetrant, the *bone lesions* "skipped" in 4 of 7 instances, a penetrance rate of 43%. Five additional "chains" of three generations of hypophosphatemics were provided by the "B", "H" and "L" kindreds, giving a total of 12. In 9 of 12 such "chains" the intervening transmitter did not have evidence of bone disease, penetrance for bone disease thus appearing to be about 25%. If penetrance

for bone lesions was calculated separately by the sex of the "middle" parent, it would appear to have been 100% for males (2 in 2) and 10% for females (1 in 10).

Penetrance for hypophosphatemia was examined in two ways. The living parents of hypophosphatemic persons were examined, and one of them invariably proved to be hypophosphatemic. This suggested that hypophosphatemia is fully penetrant. When, however, the classified children of affected persons were tabulated, hypophosphatemia appeared highly but not completely penetrant. The observation suggesting some reduction in penetrance was a significant deficiency of affected and an excess of normal daughters (but not sons) among the progeny of hypophosphatemic women.

The discrepancy can be seen best in Line 12 of Tables 1 and 2. If hypophosphatemia is inherited as a fully penetrant sex-linked dominant, affected females should have affected and normal sons and daughters in a 1:1:1:1 ratio. But the hypophosphatemic mothers in Table 1 (all of whose children were studied) form the ratio, 14:10:9:20, and the larger group of hypophosphatemic mothers in Table 2 form the ratio 20:13:10:25. The heterogeneity χ^2 with 1 d.f. (sum of χ^2 for each sex separately minus total χ^2) for the completely studied sibships of Table 1 was 3.914, $P < 5\%$. For the larger group in Table 2, heterogeneity χ^2 was 6.43, significant at the 2% point.

The total male:total female ratios 24:29 (Line 12 Table 1) and 33:35 (Table 2) are clearly not significant deviations from unity. This suggests that the deficiency of affected females did not result from loss through decreased viability. The disturbance can be seen to have resulted largely from an abnormal ratio of affected:normal females. In the completely studied sibships of Table 1 the ratio of affected:normal females is 9:20 ($\chi^2 = 4.172$, $P < 5\%$) while with all classified females, Table 2, it is 10:25 ($\chi^2 = 6.971$, $P < 1\%$).

Looking further into reasons for the excess of normal daughters from affected mothers, the deficit appears to have been produced by misclassification of young girls. Figure 6 shows that the largest disproportion between normal and affected daughters occurs in the first decade. Table F and Figure 2 show four young daughters of affected mothers in the "H" kindred whose serum phosphorus values are low with 95% confidence but not 99% confidence. Transfer of this group from the normal into the hypophosphatemic category would abolish the discrepancy and make it appear that hypophosphatemia were fully penetrant.

Sex Difference in Expression of the Abnormal Gene: It was pointed out in connection with the "E" kindred that the abnormal gene had a more striking effect in males than in females. This sex difference was manifested in three ways: 1. By a larger *proportion* of hypophosphatemic males than females who showed clinical bone disease, 2. By bone lesions of *greater severity* in males with bone disease than females with bone disease, and 3. By slightly, but significantly, lower serum phosphorus values in hypophosphatemic males than hypophosphatemic females, the sex difference in phosphorus levels being greatest between boys and girls under 16 years.

The hypophosphatemic persons are listed in Table 3, by sexes and by presence or absence of rachitic bone disease. It is apparent that most hypophosphatemic males have some evidence of bone disease (28 in 30 = 93%) while only a minority of hypophosphatemic females do (12 in 48 = 25%).

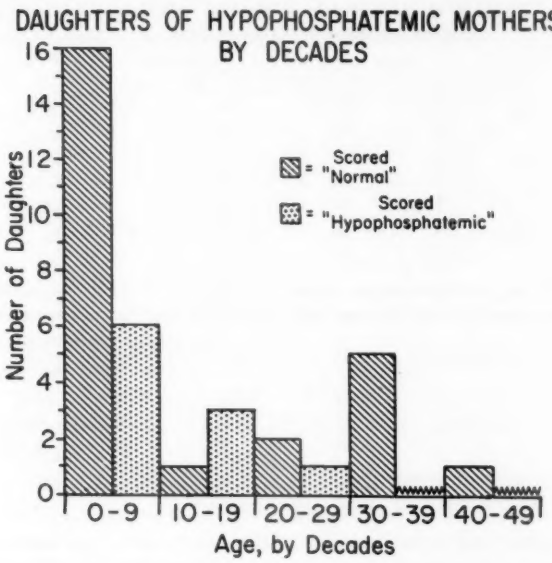

FIG. 6. A frequency distribution, by decades, of all normal and hypophosphatemic daughters of hypophosphatemic mothers from "E," "B," "H" and "L" kindreds.

The sex difference in severity of bone involvement is greater than the rough four-fold difference in thresholds suggested by the ratio 93:25. Males with bone lesions generally have severe, unequivocal disease while females scored as having bone disease rarely have even moderately severe lesions.

The third sex difference between the hypophosphatemic persons in the "E" kindred was a statistically significant difference between mean levels of serum phosphorus. The greatest mean difference was found between 6 boys and 5 girls and averaged 0.47 mg.%. The three new kindreds afford the opportunity of retesting this finding with larger groups. By pooling data on all hypophosphatemic children of the four kindreds, groups of 12 boys and 13 girls under 16 years of age can be compared. Their serum phosphorus values were adjusted to age 2, separately by sexes, using the normal regression curves of age on phosphorus published earlier (14). Means and standard errors of these adjusted values were computed and are shown in Table 4. It will be seen that the difference in means is 0.46 mg.%, essentially the same as for the "E" kindred alone. The standard errors are small, and the "t" test is highly significant. This test suggests that hypophosphatemic boys hemizygous for the abnormal gene have significantly lower serum phosphorus values than hypophosphatemic girls who are heterozygous.

A mathematically more complex, but possibly less sensitive, test of this sex difference was performed for us by Professor Bernard Greenberg. He used for his test portions of the regression curves fitted to the *abnormal* populations separately by

TABLE 3. HYPOPHOSPHATEMIC PERSONS WITH AND WITHOUT CLINICAL EVIDENCE OF BONE DISEASE, BY SEXES

Kindred	Males		Females		Total
	With bone disease	Without bone disease	With bone disease	Without bone disease	
"E"	13	2	6	16	37
"B"	6	0	5	9	20
"H"	6	0	0	7	13
"L"	3	0	1	4	8
Totals	28	2	12	36	78

TABLE 4. AGE ADJUSTED SERUM PHOSPHORUS VALUES OF 12 HYPOPHOSPHATEMIC BOYS AND 13 HYPOPHOSPHATEMIC GIRLS LESS THAN 16 YEARS OF AGE

Sex	Average Age	No.	d.f.	Serum Phosphorus, mg.%		
				Mean, mg.%	Std. error	Sum of squares
Boys	5.5 yrs.	12	11	2.92	.082	.8968
Girls	6.6 yrs.	13	12	3.38	.093	1.3572
Total		25	23	.46		2.2540

difference of means = .46 mg.%; $t = 3.67$, 23 d.f., $P < .01$

sexes and described elsewhere (14). The difference between the curves for affected boys and affected girls proved not to be statistically significant.

Linkage

The four kindreds were examined systematically for presence of colorblindness. Only one colorblind person was found, a woman who had recently married into the family. Although no tests were made for any of the rarer sex-linked conditions, there is no reason to think that any were segregating in the kindreds. For example, no "bleeders" or familial eye diseases were reported or encountered.

Environmental Factors Influencing Expression of Hypophosphatemia

In the report on the "E" kindred it was pointed out that the members of the family had no peculiar dietary habits which might lead to rickets. On the contrary, most of them were out-of-doors workers eating diets rich in dairy products. The dietary histories of the members of the "B", "H" and "L" kindreds were similar to those of the "E" kindred, and most of the persons had occupations which kept them out-of-doors much of the time.

Maternal age and parity did not appear to affect the expression of the abnormal gene in the original "E" kindred. An effect was sought in the combined kindreds by the method of Haldane and Smith (6). The raw birth rank data from the three new kindreds are listed in Table J. The data from the 15 completely studied sibships of the "E" kindred, plus 8 completely studied new sibships from the "B", "H" and "L" kindreds, plus 3 sibships in which the missing information concerned only first or last children (2 from "E", 1 from "H" kindred) are combined for the test shown in Table 5. The fact that expected and observed mean birth ranks (6A and M)

TABLE 5. BIRTH RANK TEST (HALDANE AND SMITH)

Sibship Size	Number of Sibships	Birth Rank of Affected Persons							Total Affected	Total Siblings
		1	2	3	4	5	6	7		
2	11	5	6						11	22
3	7	2	4	5					11	21
4	4	2	0	5	2				9	16
5	3	1	1	1	1	2			6	15
6	0	0	0	0	0	0	0		0	0
7	1	1	0	0	0	1	1	0	3	7
Totals	26	11	11	11	3	3	1	0	40	81

6A = 594, M = 558, V = 1128, S = 33.59

differ by less than two standard deviations (2S) means that there was not a significant birth rank effect in these 26 sibships.

DISCUSSION

The material contained in this paper confirms all but one of the conclusions reached in our study of the "E" kindred. The "B", "H" and "L" kindreds add 29 new sibships containing 15 affected males and 26 affected females to the literature. We have, therefore, diagnosed in North Carolina a total of 78 cases of familial hypophosphatemia in a population of 4.5 million persons. This does not represent complete ascertainment, because others have reported similar families (9). We suggest that sex-linked dominant familial hypophosphatemia should not be considered a rare and exotic entity. This concentration of cases in a limited geographical area raises the question whether the four kindreds are related. Three of the kindreds ("E", "B" and "L") live within a relatively small sector of the Coastal Plain of North Carolina. The fourth kindred ("H") is largely located in a community 100 miles to the west. No connection as recent as the last 6 generations has been established between any of the kindreds. Since the state was settled by emigration largely from the British Isles 200–250 years ago, however, it is quite possible that such a relationship exists.

In all three of the new kindreds, hypophosphatemia is inherited as a sex-linked dominant trait. Seven additional affected fathers had had 17 children, 15 of whom could be classified. These children consisted of 6 normal sons and 9 affected daughters. Thus, in our 4 kindreds, 14 affected fathers have produced 37 children, 16 normal sons and 21 affected daughters, without a single instance of father-to-son transmission of hypophosphatemia or a single normophosphatemic daughter. The possibility is very remote that this distribution is merely an abnormal sample from a population segregating for an autosomal dominant trait.

The characteristics of the homozygous state still remain unknown. We have not yet encountered a marriage of two affected persons, apparently because the rate of inbreeding has been low in the native white people of the Coastal Plain.

The new kindreds, like the original one, demonstrate that penetrance for hypophosphatemia is high in both sexes. A new observation, however, is that penetrance for hypophosphatemia is not *complete* in females (see below). Penetrance for overt bone disease is clearly high among males but low among females. Among mothers

with affected fathers and affected children, penetrance for the bone disease is only 10%. If hypophosphatemic females of all ages are considered, however, bone lesions can be detected in about 25%.

The three new kindreds showed again the sex difference with respect to severity of bone lesions. The bone lesions of the males in the "E", "H" and "L" kindreds were more severe than the bone lesions of the females. The "B" kindred is exceptional, because the proband was a woman, and the bone lesions in her and her sisters were definitely more severe than those of the hypophosphatemic women of the other kindreds. Reference to Table A, will emphasize this. All of these women, IV-44 through IV-48, had had osteotomies between ages 3 and 6. Furthermore, 3 of the 5 women had had at least one Caesarian section because of rachitic pelvic deformity. The lesions in these women appear to be as severe as those of the males in this and the other kindreds. Yet these women were not homozygous, because their mother was normophosphatemic as were some of their children.

Another finding in the "E" kindred had been a significantly lower serum inorganic phosphorus level in affected males than affected females. The greatest difference had been found between 6 affected boys and 5 affected girls and had amounted to 0.47 mg.%. The new kindreds provided 6 additional affected boys and 8 affected girls. When these were added to the previous examples, giving groups of 12 and 13, essentially the same mean difference was observed (0.46 mg.%), the males again being lower. The means of the two groups were found to be significantly different when tested in the manner previously described. When the sex difference was tested by comparing the curvilinear functions fitted separately by sexes to the abnormal populations, however, the difference was not significant. The last result may imply that there is no real sex difference for serum inorganic phosphorus among affected persons. On the other hand, it may mean that the complex functions fitted to the small abnormal populations had excessively large variances. If a real difference in serum phosphorus exists between hypophosphatemic boys and girls, it is slight. One is tempted to suggest that the accompanying normal allele in the heterozygous female accounts for the difference. How this might operate is not clear. We were unable, for example, in the "E" kindred to show any difference in phosphate re-absorption in the kidneys of affected males and females (14).

The last confirmation of our original findings was with respect to lack of an effect of maternal age or parity on expression of the gene. The Haldane and Smith test applied to 40 affected persons among a total of 81 in 26 sibships showed no effect of birth rank (Table 5).

A new and at first puzzling finding was the distribution of children of affected females (Line 12, Tables 1 and 2). There was a slight but not significant excess of affected sons and a significant deficiency of affected daughters. The reason for the discrepancy with the daughters can be clearly deduced from the operating characteristics of the discriminant devised for defining the normal limits for serum phosphorus (5). We pointed out in this connection that the discriminant is highly efficient in correctly categorizing the hypophosphatemics, both males and females, i.e. not more than 0.5% of the normals will be incorrectly diagnosed as abnormals. Furthermore, it was pointed out that while an *average* of not more than about 6% of abnormal

persons should be incorrectly diagnosed as normal, the rate of false negatives is expected to be higher among young females. It is among the young "normal" daughters, therefore, that most of the incorrectly classified persons are to be expected. We are of the opinion that this diagnostic error has occurred and explains the apparent excess of normal daughters from affected mothers. It is presumed that the four young girls of the "H" kindred whose phosphorus values are low with 95% but not 99% confidence will be scored as abnormal when they become older. We anticipate that their phosphorus values will slip below the confidence limit as its discriminatory power becomes greater (lower rate of false negatives at older ages).

It is of some interest that the presumed false negative diagnoses (lack of penetrance) involves daughters of affected mothers but not daughters of affected fathers. All 21 daughters of affected fathers who were studied were clearly abnormal. The low rate of false negatives in this circumstance may be related to the fact that 17 of the 21 daughters of affected fathers were more than 15 years of age. It is also possible that the study was unconsciously biased in another manner, i.e. that we examined the daughters of affected fathers more carefully because of their importance to the genetic argument.

Finally, it is interesting to note how an arbitrarily set serum phosphorus discriminant produces a reduction in penetrance among affected females. The phosphorus discriminant was deliberately set at the 99% point to minimize false positive diagnoses. This in turn increased the probable number of false negatives. The ratios can be shown to come into balance if the 95% confidence limit rather than the 99% confidence limit is used. It will be very interesting to follow the girls thought to have been incorrectly scored as normal to see whether their phosphorus values become clearly abnormal as they become older. If this does not occur and they can be shown to be genetically abnormal through their progeny, it may be wise to adopt the 95% confidence limit as the discriminant. It is for this reason that both limits will be published in our definitive paper on the normal range of serum phosphorus (5).

SUMMARY

1. Three additional kindreds of sex-linked dominant familial hypophosphatemia and vitamin D resistant rickets including 41 new patients are described.

2. It is shown that the new families conform in most respects to the previously described "E" kindred, i.e.

a) Hypophosphatemia appears highly penetrant. Only in young girls does there appear to be any reduction in penetrance.

b) Penetrance for bone lesions is clearly and greatly reduced in females.

c) In hypophosphatemic mothers, penetrance for bone lesions is about 10%. Some types of bone lesions can be found, however, in 25% of hypophosphatemic females of all ages.

d) Bone lesions are more severe in males than in females.

e) Hypophosphatemic girls appear to have a slightly higher level of serum inorganic phosphorus than hypophosphatemic boys.

f) Maternal age and parity appear *not* to affect expression of the gene.

3. The observation of an excess of apparently normal females from affected mothers

(incomplete penetrance for hypophosphatemia in females) is probably an error in falsely diagnosing genetically affected girls as normal. This error is an empirical confirmation of the expected higher rate of false negative diagnoses in young females than older females or males, an expectation previously arrived at on theoretical grounds.

REFERENCES

1. Albright, F., Butler, A. M., and Bloomberg, E. 1937. Rickets resistant to vitamin D therapy. *Am. J. Dis. Child.* 54: 529–547.
2. Albright, F., Consalazio, W. V., Coombs, F. S., Sulkowitch, H. W., and Talbott, J. H. 1940. Metabolic studies and therapy in a case of nephrocalcinosis with rickets and dwarfism. *Bull. Johns Hopkins Hosp.* 66: 7–33.
3. Fanconi, G. 1936. Der frühinfantile nephrotische-glykosurische Zwergwuchs mit hypophosphatämischer Rachitis. *Jahrb. Kinderh.* 147: 299–338.
4. Gerstenberger, H. J. 1933. Rachitis hepatica. *Monatsschr. Kinderh.* 56: 217–226.
5. Greenberg, B. G., Winters, R. W., and Graham, J. B. The normal range of serum inorganic phosphorus and its use as a discriminant in the diagnosis of congenital hypophosphatemia. *J. Clin. Endocr. and Metab.* (in press).
6. Haldane, J. B. S., and Smith, C. A. B. 1948. A simple exact test for birth-order effect. *Ann. Eugen.* 14: 117–124.
7. Lowe, C. V., Terrey, M., and MacLachlan, E. A. 1952. Organic aciduria, decreased renal ammonia production, hydrophthalmos, and mental retardation. A clinical entity. *Am. J. Dis. Child.* 83: 164–184.
8. McCance, R. A. 1947. Osteomalacia with Looser's nodes (Milkman's Syndrome) due to a raised resistance to vitamin D acquired about the age of 15 years. *Q. J. Med.*, 16: 33–45.
9. Mitchell, F. N., and Mitchell, J. E. 1957. Vitamin D-resistant rickets. *Am. J. Dis. Child.* 93: 385–390.
10. Parsons, L. G. 1927. The bone changes occurring in renal and coeliac infantilism, and their relationship to rickets. Part I. Renal rickets. *Arch. Dis. Child.* 2: 1–25. Part II. Coeliac rickets. *Arch. Dis. Child.* 2: 198–211.
11. Rathbun, J. C. 1948. 'Hypophosphatasia', a new developmental anomaly. *Am. J. Dis. Child* 75: 822–831.
12. Swann, G. F. 1954. Pathogenesis of bone lesions in neurofibromatosis. *Brit. J. Radiol.* 27: 623–629.
13. Winters, R. W., Graham, J. B., Williams, T. F., McFalls, V. W., and Burnett, C. H. 1957. A genetic study of familial hypophosphatemia and vitamin D resistant rickets. *Tr. Ass. Am. Physicians* 70: 234–242.
14. Winters, R. W., Graham, J. B., Williams, T. F., McFalls, V. W., and Burnett, C. H. 1958. A genetic study of familial hypophosphatemia and vitamin D resistant rickets with a review of the literature. *Medicine* 37: 97–142.
15. Winters, R. W., McFalls, V. W., and Graham, J. B. 1960. A girl with sporadic hypophosphatemia and vitamin D resistant rickets. *Pediatrics* (in press).

APPENDIX

TABLE A. CLINICAL DATA ON HYPOPHOSPHATEMIC MEMBERS OF "B" KINDRED

Generation and Number	Age When Studied	Sex	History	Height[1]	Physical and Radiologic Findings[2]
			Children		
V-1	14	M (Male)	+	25	Moderately severe rickets during early childhood, successful osteotomy at age 6 years, slight residual *genu varum* deformity.
V-18	4	M	+	10	Active rickets of moderate severity with slight *genu varum*.
V-81	3	M	+	<3	Severe active rickets, partly treated; moderate *genu varum*.
V-82	11	M	+	<3	Active rickets of moderate severity with moderate *genu varum*.
V-2	11	F (Female)	−	10	Negative
V-66	9 mos.	F	−	25	Negative
V-83	12	F	−	50	Negative
V-87	3	F	−	25	Negative
			Adults		
III-1	60	M	+	62	Severe *genu varum*; moderate anterior bowing of tibiae. Degenerative joint disease at knees.
III-10	56	M	+	63	Moderate *genu varum*; degenerative joint disease of knees.
IV-1	37	F	−	59	Negative
IV-7	26	F	−	58	Negative
IV-8	24	F	−	59	Negative
IV-37	26	F	−	63	Negative
IV-38	31	F	−	60	Negative
IV-44	18	F	+	61	Moderately severe deformities during childhood; successful corrective osteotomy at age 6 years. Slight residual *genu varum* deformity.
IV-45 (Proband)	25	F	+	57	Moderately severe rickets during childhood;[3] successful corrective osteotomy at age 5 years. Moderate residual *genu varum* deformity. Caesarian section both pregnancies.
IV-46	27	F	+	58	Active rickets of moderate severity during childhood;[3] successful osteotomy at age 3 years. Slight residual *genu varum* deformity. Caesarian section one time.
IV-47	34	F	+	59	Active rickets of moderate severity during childhood,[3] successful osteotomy at age 4 years. Slight residual *genu varum* deformity. Caesarian section three times.
IV-48	30	F	+	58	Active rickets of moderate severity during childhood;[3] successful osteotomy at age 3 years. Slight residual *genu varum* deformity.

[1] Height of children in percentiles; height of adults in inches.
[2] Based upon physical examination and x-rays of long bones.
[3] Verified by old roentgenograms taken at the time of osteotomies.

TABLE B. BLOOD CHEMICAL DATA ON HYPOPHOSPHATEMIC MEMBERS OF "B" KINDRED*

Generation and Number	Age When Studied	Year of Birth	Sex	Ca^{++} mg %	T.P. gm %	P mg %	Alkaline Phosphatase B.U.	Na^{+} mEq/L	K^{+} mEq/L	Cl^{-} mEq/L	BUN mg %
Children											
V-1	14	1943	M (Male)	9.1	6.8	2.8	17.4	137	5.0	103	9
V-18	4	1953	M	10.5	6.5	2.4	16.5	—	—	—	—
V-81	3	1955	M	10.3	6.7	3.1	11.1	136	5.4	102	10
V-82	11	1946	M	9.5	7.0	2.4	11.3	141	3.8	103	16
V-2	11	1946	F (Female)	9.8	6.3	3.0	14.1	—	—	—	—
V-66	9 mos.	1957	F	—	—	3.2	17.3	—	—	—	—
V-83	12	1945	F	9.3	6.7	3.0	5.7	139	4.2	106	13
V-87	3	1955	F	—	—	3.2	—	—	—	—	—
Adults											
III-1	60	1897	M	9.0	6.8	2.0	4.4	139	4.3	103	14
III-10	56	1901	M	9.3	7.1	2.1	5.1	141	3.8	103	11
IV-1	37	1920	F	9.8	6.6	1.8	3.9	138	4.2	103	8
IV-7	26	1931	F	9.2	6.5	2.4	4.4	140	3.9	105	10
IV-8	24	1933	F	9.2	7.1	2.2	5.3	142	4.4	106	11
IV-37	26	1931	F	9.2	7.0	2.0	5.7	139	4.0	105	16
IV-38	31	1927	F	—	—	1.9	—	—	—	—	—
IV-44	18	1939	F	9.8	7.4	2.7	4.3	137	4.6	104	12
IV-45 (Proband)	25	1932	F	9.0	7.1	2.6	4.4	138	3.9	101	11
IV-46	27	1930	F	9.5	6.9	2.4	3.6	140	4.3	104	11
IV-47	34	1923	F	9.8	6.8	2.0	5.1	140	4.5	103	10
IV-48	30	1928	F	9.9	7.0	2.4	—	142	4.3	105	12

* T.P. = total protein; P = serum inorganic phosphorus; BUN = blood urea nitrogen.

TABLE C. SERUM PHOSPHORUS LEVELS OF NORMAL MEMBERS OF "B" KINDRED*

Generation and Number	Sex	Age When Studied	Year of Birth	P (mg %)
Children				
IV-42	M (Male)	7	1950	4.1
V-69	M	4	1953	5.3
V-70	M	10	1947	4.8
V-71	M	8	1949	4.6
V-79	M	8	1949	4.3
V-85	M	9	1948	4.6
V-19	F (Female)	4	1953	5.5
V-20	F	5	1952	5.2
V-21	F	1	1956	5.4
V-65	F	6	1951	4.6
V-72	F	11	1946	5.3
V-80	F	8	1949	4.3

TABLE C (*Continued*)

Generation and Number	Sex	Age When Studied	Year of Birth	P (mg%)
		Adults		
III-4	M	66	1891	3.1
III-8	M	64	1892	2.9
III-9	M	60	1897	2.9
III-18	M	56	1901	2.8
III-19	M	57	1900	4.0
III-20	M	42	1915	2.7
III-21	M	49	1908	3.5
IV-2	M	35	1922	3.3
IV-3	M	33	1924	3.9
IV-4	M	28	1929	4.1
IV-5	M	25	1932	3.7
IV-9	M	27	1930	2.9
IV-10	M	22	1935	4.2
IV-39	M	38	1919	4.3
IV-40	M	36	1921	3.4
IV-41	M	29	1928	4.0
V-3	M	17	1940	3.5
II-4	F	74	1883	4.0
III-12	F	57	1900	4.0
III-22	F	53	1904	4.0
V-84	F	16	1941	3.5

* P = serum inorganic phosphorus.

TABLE D. CLINICAL DATA ON HYPOPHOSPHATEMIC MEMBERS OF "H" KINDRED

Generation and Number	Age When Studied	Year of Birth	Sex	History	Height[1]	Physical Findings[2]
			Children			
IV-2 (Proband)	4	1954	M (Male)	+	<3	Moderately severe active rickets with *genu varum*; slight anterior gowing of tibiae; pigeon breast.
V-8	6	1951	M	+	25	Moderate *genu varum*
V-6	5	1952	F (Female)	−	25	Negative
			Adults			
III-11	48	1909	M	+	60	Marked *genu varum*
III-12	41	1916	M	+	60	Moderate *genu varum*
III-13	33	1924	M	+	62	Moderate *genu varum*
III-25	50	1907	M	+	59	Moderate *genu varum*
II-3	67	1890	F	−	62	Negative
III-1	32	1925	F	−	63	Negative
III-8	24	1933	F	−	60	Negative
III-9	22	1935	F	−	61	Negative
IV-35	26	1931	F	−	61	Negative
IV-38	21	1936	F	−	59	Negative

[1] Height of children as percentile; height of adults in inches.

[2] Based on physical examination.

TABLE E. BLOOD CHEMICAL DATA ON HYPOPHOSPHATEMIC MEMBERS OF "H" KINDRED*

Generation and Number	Age When Studied	Year of Birth	Sex	Ca^{++} mg %	T.P. gm %	P mg %	Alkaline Phosphatase B.U.	Na^{+} mEq/L	K^{+} mEq/L	Cl^{-} mEq/L	BUN mg %
Children											
IV-2 (Proband)	3	1954	M (Male)	9.7	6.1	2.7	12.9	138	4.2	102	9
V-8	6	1951	M	11.1	7.4	2.9	16.8	139	3.8	105	14
V-6	5	1952	F (Female)	9.1	6.8	3.0	9.4	141	3.9	106	—
Adults											
III-11	48	1909	M	10.9	6.9	2.2	2.4	141	4.1	105	—
III-12	41	1916	M	11.0	7.0	2.2	3.1	141	4.0	100	—
III-13	33	1924	M	10.1	6.9	2.4	4.4	142	3.8	108	19
III-25	50	1907	M	10.7	6.6	2.2	2.3	144	4.8	107	17
II-3	67	1890	F	10.9	7.2	2.6	4.5	142	4.3	104	—
III-1	32	1925	F	—	—	2.4	2.8	—	—	—	—
III-8	24	1933	F	—	—	2.4	—	—	—	—	—
III-9	22	1935	F	—	—	2.3	—	—	—	—	—
IV-35	26	1931	F	11.6	8.4	2.5	2.3	137	4.7	103	17
IV-38	21	1936	F	10.8	7.0	2.7	4.1	141	3.7	107	6

* T.P. = total protein; P = serum inorganic phosphorus; BUN = blood urea nitrogen.

TABLE F. BLOOD CHEMICAL DATA ON NORMAL MEMBERS "H" KINDRED[1]

Generation and Number	Sex	Age When Studied	Year of Birth	P (mg %)
		Children		
IV-1	M (Male)	7	1950	5.0
IV-19	M	6	1951	4.7
IV-36	M	9	1948	4.5
V-3	M	8	1949	4.3
V-4	F (Female)	6* (7)†	1952	4.3* (3.7)†
V-5	F	6* (7)†	1952	4.4* (3.7)†
V-7	F	2* (3)†	1956	5.5* (3.8)†
V-9	F	4* (5)†	1953	5.2* (3.9)†
V-10	F	2†	1957	5.2†
		Adults		
II-7	M	64	1893	4.4
III-4	M	30	1927	4.0
III-5	M	28	1929	4.3
III-6	M	26	1930	4.1
III-7	M	22	1935	4.2

TABLE F (*Continued*)

Generation and Number	Sex	Age When Studied	Year of Birth	P (mg %)
		Adults		
III-14	M	45	1911	3.4
III-15	M	24	1933	3.8
III-18	F	43	1914	3.8
III-19	F	36	1921	3.4
III-20	F	31	1926	3.7
III-31	F	64	1893	4.5

* Determinations in 1958, not fasting. Clearly normal.

† Determinations in 1959, fasting. The values within parentheses are significantly low at the 95% point, but not at 99% point.

[1] *P* = serum inorganic phosphorus.

TABLE G. CLINICAL DATA ON HYPOPHOSPHATEMIC MEMBERS OF "L" KINDRED

Generation and Number	Age When Studied	Sex	History	Height[1]	Physical Findings[2]
			Children		
V-2	15	F (Female)	–	59	Negative
V-6	4	F	–	25	Negative
			Adults		
IV-1 (Proband)	37	M (Male)	+	63	Moderate *genu varum*
IV-2	35	M	+	62	Moderate *genu varum*
IV-6	23	M	+	64	Moderate *genu varum*
III-4	64	F	–	63	Negative
III-5	51	F	+	60	Moderate *genu varum*
V-1	16	F	–	61	Negative

[1] Height of children in percentiles. Height of adults in inches.

[2] Based on physical examination.

TABLE H. BLOOD CHEMICAL DATA ON HYPOPHOSPHATEMIC MEMBERS OF "L" KINDRED*

Generation and Number	Age When Studied	Year of Birth	Sex	Ca^{++} (mg %)	T.P. (gm %)	P (mg %)
			Children			
V-2	15	1942	F (Female)	—	—	2.8
V-6	4	1952	F	—	—	3.1
			Adults			
IV-1 (Proband)	37	1919	M (Male)	9.8	6.9	2.4
IV-2	35	1921	M	9.6	6.7	2.3
IV-6	23	1933	M	10.3	7.0	2.5
HI-4	64	1891	F	—	—	1.7
III-5	51	1905	F	—	—	2.5
V-1	16	1939	F	—	—	2.6

* T.P. = total protein; P = serum inorganic phosphorus.

TABLE I. BLOOD CHEMICAL DATA ON NORMAL MEMBERS OF "L" KINDRED*

Generation and Number	Sex	Age When Studied	Year of Birth	P (mg %)
		Children		
IV-9	M (Male)	9	1946	3.4
V-7	M	12	1944	3.8
V-8	M	10	1946	4.4
IV-8	F (Female)	15	1941	4.0
V-4	F	7	1948	4.8
V-5	F	5	1951	4.9
		Adults		
III-11	M	46	1909	4.4
III-10	F	51	1904	4.5
IV-3	F	31	1925	3.8
IV-4	F	29	1927	4.5
IV-7	F	31	1924	3.4

* P = serum inorganic phosphorus.

TABLE J. BIRTH ORDER OF CLASSIFIED CHILDREN OF HYPOPHOSPHATEMIC PARENTS, OF MEMBERS OF SIBSHIPS SEGREGATING FOR HYPOPHOSPHATEMIA OR BOTH

"B" Kindred			"H" Kindred			"L" Kindred		
Sibship size	Affected parent	Birth order*	Sibship size	Affected parent	Birth order*	Sibship size	Affected parent	Birth order*
			Phenotype of Parent Established—*All* Progeny Classified					
1	Mother	N	1	Father	N	1	Father	A
1	Mother	A	2	Mother	NA	1	Father	A
2	Mother	NN	3	Mother	ANN	1	Father	A
2	Mother	NA	5	Mother	N(NN)AN			
3	Mother	NAA						
3	Mother	NNA						
3	Mother	AAN						
2	Father	AA						
			Phenotype of Parent Established—*Not All* Progeny Classified					
2	Mother	--	4	Father	AAN-	6	Mother	AANN-A
3	Mother	N-A	10	Mother	ANNA-NAN-N			
7	Father	ANN-NAN						
			Phenotype of Parent *Not* Established—*All* Progeny Classified					
3	—	NAN						
			Phenotype of Parent *Not* Established—*Not All* Progeny Classified					
4	—	----	7	—	A--N---	5	—	A--A-
7	—	-------	8	—	---A--N-			
10	—	NNAANAA-AN	10	—	AAN-NAN-N-			

* A = affected, i.e. hypophosphatemic; N = normophosphatemic.

Hereditary Chronic Kidney Disease: An Alternative to Partial Sex-Linkage in the Utah Kindred

JOHN B. GRAHAM[1]

Department of Pathology, University of North Carolina

A LARGE KINDRED residing in Utah and containing many persons with chronic kidney disease was first described in 1951 (7, 11). The unusual distribution of affected persons appeared consistent with dominant partial sex-linkage, a mode of inheritance postulated to occur in humans by Haldane (4). This kindred has been studied again after a 5 year interval, and the second report has recently been published (8). The authors re-affirm their earlier belief that they are dealing with an example of dominant partial sex-linkage.

The kindred, as originally described, was analyzed by the method of sequential analysis by Morton (6), who was unable to reject partial sex-linkage in this kindred with the same assurance as in Haldane's original group of diseases. However, he obviously felt strongly that the disorder was inherited in some other fashion, because he suggested an alternative which is highly unlikely on *a priori* grounds, i. e. dominant complete sex-linkage for most of the kindred and some other disease altogether to account for the several examples of father-son transmission. As my title suggests, I agree with Morton that all possible alternatives should be excluded before partial sex-linkage is accepted, because, like him, I doubt that partial sex-linkage occurs. In the course of examining the alternatives, I believe that I have discovered the correct mode of inheritance. While my explanation is somewhat complex, the *a priori* probability that it is correct is probably greater than that of Morton's suggestion, and it is, like his, not impossible on cytological grounds.

The alternative which I am proposing is attractive, because it explains certain features of the kindred which partial sex-linkage will not explain. Specifically, the data are consistent with an autosomal, dominant trait, more severe in males. There is a deficiency of males over the entire kindred, due to a deficiency of affected males among the children of affected parents of *both* sexes. This deficiency results, presumably, from a significantly higher death rate of affected males very early during intrauterine life, because the kindred as reported does not mention frequent abortions, miscarriages or neonatal deaths. The fact that the mutant gene is pleiotropic and incompletely penetrant in addition to having an expression which varies with the sex of the affected person makes it easy to understand why partial sex-linkage has been mistakenly proposed.

Received Feb. 25, 1959.

[1] The investigations of the author have been supported by research grants H-1333 and H-3140 of the National Heart Institute, Public Health Service.

INHERITED CHRONIC RENAL DISEASE IN OTHER KINDREDS

A considerable amount of literature on inherited chronic "Bright's disease" has accumulated since about 1875. The earlier reports are listed in the bibliographies of Alport (1) and Eason and others (2). These earlier kindreds will not be summarized, because important persons in the kindreds were not studied and methods of study varied widely. When the kindreds reported by Alport (1), Reyersbach and Butler (9), Sturtz and Burke (12), Hamburger and others (5), Goldbloom and others (3), and Robin and others (10), are collated, however, certain recurring features become apparent. These may be summarized as follows:

1. Inherited chronic renal disease has usually been associated with nerve deafness, occasionally with eye lesions.

2. The renal disease has been more severe in males than females; the males usually have died young without issue, while the females often have lived to a ripe age and have had many children.

3. Few females with renal disease have also been deaf, but many males have shown both defects.

4. A few persons have shown only deafness.

5. Rarely, persons who have appeared completely normal, and occasionally persons who have shown only deafness, have transmitted the kidney disease.

6. Male-to-male transmission of the disorder has been observed, but only in Robin's kindred (only one other of the 51 affected males in this group of kindreds had a son, and he was normal).

7. The severity of the disorder has varied from kindred to kindred (see Robin).

8. The clinical and pathological picture in some kindreds has appeared to be

TABLE 1. OFFSPRING OF AFFECTED PERSONS (FROM LITERATURE, EXCLUSIVE OF KINDRED REPORTED BY PERKOFF AND OTHERS)

Source of data	Affected Males							Affected Females						
				Sons		Daughters					Sons		Daughters	
	Total males	No. deaf	No. of fathers	Affected	Normal or unknown	Affected	Normal or unknown	Total females	No. deaf	No. of mothers	Affected	Normal or unknown	Affected	Normal or unknown
Alport	7	3	0	0	0	0	0	12	5	6	7	4	10	5
Reyersbach	15	12	1	0	1	0	0	9	1	8	15	8	8	14
Sturtz	8	8	0	0	0	0	0	5	1	5	4	7	3	7
Hamburger	7	4	0	0	0	0	0	6	0	3	7	3	5	3
Goldbloom	3	2	0	0	0	0	0	1	0	1	3	1	0	0
Robin	11	?	6	3	4	5	4	6	?	1	3	3	1	0
Totals	51	>29	7	3	5	5	4	39	>7	24	39	26	27	29
Sex ratios				8		9					65		56	
Total children				17							121			

that of "chronic pyelonephritis" while in others it has appeared to be "chronic glomerulonephritis". Nerve deafness has been observed in persons with both "types" of chronic renal disease.

Table 1 shows the marked difference in fitness between affected males and affected females in these kindreds. It will be noted that only 7 of 51 affected males had children, producing only 17. On the other hand 24 of 39 affected women had children, a total of 121. Also, Table 1 demonstrates that deafness has been more common among affected males (29/51) than females (7/39). Finally, it should be noted that the various ratios are not markedly abnormal among the progeny of affected persons, although there has been an excess of affected children and also an excess of males.

ANALYSIS OF THE RE-EXAMINED UTAH KINDRED

The re-evaluation of the Utah kindred (8) is notable in several respects. This kindred is far and away the largest yet reported, and it has been studied very extensively. Considerably more people have been examined (168) than 5 years ago, and 97 persons have been re-examined. Also the kidneys of 4 living patients and 1 dead patient have been biopsied, and another affected person has been autopsied.

It is interesting to note that the kidneys at the most recent autopsy suggested "chronic glomerulonephritis", because the published pictures from an earlier autopsy clearly showed the changes of "chronic pyelonephritis" (7). The occurrence of both forms of chronic renal disease in a single kindred also having nerve deafness is an interesting observation, because in each of the kindreds reported earlier nerve deafness had been associated with only one type of chronic renal disease.

The sex-limited nature of the deafness is shown in Table 2. Here it can be seen that 19 of 20 cases of deafness were in males, and that deafness was inherited by these men from affected parents of both sexes. Furthermore, deafness is associated with renal disease too frequently for this combination to represent the coincidence of independent double heterozygosity. It is more reasonable to assume that deafness and renal disease are manifestations of the same pleiotropic gene. The additional evidence for pleiotropy is that some persons who are deaf alone have transmitted renal disease (III-7, III-43), and some who have had only renal disease have transmitted deafness (II-16, II-17, III-1, III-36).

The really peculiar nature of the Utah kindred begins to appear in Table 3 where the sex ratios are listed by generations, all persons considered. The ratios are reversed in generations II, III and IV, where the females considerably outnumber the males. The excess of females is statistically significant when the entire kindred is considered and unity expected ($\chi^2 = 4.41$, $p < .04$).

It is important to emphasize at this point my definition of an affected person as

TABLE 2. DISTRIBUTION OF DEAFNESS AMONG AFFECTED OFFSPRING IN THE UTAH KINDRED

Transmitting Parents	Deafness Alone		Deafness Plus Kidney Disease		Totals
	Males	Females	Males	Females	
Mothers	4	0	11	1	16
Fathers	3	0	1	0	4
Totals	7	0	12	1	20

TABLE 3. SEX RATIOS AMONG OFFSPRING IN THE UTAH KINDRED, BY GENERATIONS

Generation No.	Males	Females	Unknown	Total	Ratio M:F
I	1	1	0	2	1.0
II	1	9	0	10	0.11
III	25	34	1	60	0.74
IV	86	104	0	190	0.73
V	8	8	0	16	1.0
Totals	121	156	1	278	0.78

TABLE 4. DISTRIBUTION OF GRANDCHILDREN OF THE SEVEN AFFECTED FEMALES IN GENERATION II OF THE UTAH KINDRED

Affected Parents in Gen. III	Grandsons in Gen. IV				Granddaughters in Gen.IV				Totals
	Normal	Affected	Unk.	Total	Normal	Affected	Unk.	Total	
13 Males*.....	15	5	3	23	10	21	2	33	56
13 Females†.....	16	8	5	29	22	16	5	43	72
26 Parents.....	31	13	8	52	32	37	7	76	128

* 12 of 13 males show clinical effects of abnormal gene.

† 4 of 13 females show clinical effects of abnormal gene.

this may be at variance with that of the authors and is crucial to the scoring procedure. I define an affected person as:

1. A member of the kindred who shows either deafness or renal disease or both, or
2. A member of the kindred who has had at least 1 child showing either deafness, renal disease or both.

Table 4 records the phenotypes of the grandchildren of the 9 affected women in Generation II when affected persons are scored by my criteria. The grandchildren have been tabulated separately for affected fathers and affected mothers. If complete sex-linkage were the mode of inheritance, the affected fathers of Generation III should have had only normal sons and affected daughters. Complete sex-linkage is rejected, because there was male-to-male transmission in 5 instances, and 10 normal daughters were observed. The former is the more important observation, because some of the "normal" daughters might well represent asymptomatic carriers (9 of the 10 asymptomatic carrier parents of Generation III were female).

If partial sex-linkage were the mode of inheritance (and some crossing-over had occurred), it would be expected that the affected grandchildren from affected fathers would be predominantly female, and the affected grandchildren from affected mothers would be equally divided between males and females.

Examination of the first line in Table 4 (offspring of affected males) appears at first glance to support partial sex-linkage with some crossing-over, because there is not a 1:1:1:1 ratio, but an excess of affected females and a deficiency of affected males. It should be pointed out, however, that there is also a total deficiency of males from affected fathers, the male: female ratio being 23:33.

The 2nd line of Table 4, children of affected mothers, is disturbing from the standpoint of partial sex-linkage, because here also there is an abnormal sex ratio (29M:43F) caused by a deficiency of affected males.

TABLE 5. TESTS OF SIGNIFICANCE OF RATIOS IN TABLE 4, ASSUMING INHERITANCE TO BE SIMPLE AUTOSOMAL DOMINANCE WITH FULL PENETRANCE (1:1:1:1)

Ratios Tested	Adjusted χ^2	P, 1 d.f.
Total males v. total females (52:76)	4.49	$<.04$
Affected v. normal persons (50:63)	1.51	$>.20$
Affected females v. normal females (37:32)	0.33	$>.50$
Affected males v. normal males (13:31)	7.40	$<.01$
Normal males v. normal females (31:32)	0.03	$>.80$
Affected males v. affected females (13:37)	11.52	$<.01$
Heterogeneity (1:1:1:1 ratio)		
a. Actual (31:13:32:37)	5.38	$<.03$
b. With 24 additional affected males (31:37:32:37)	0.0062	$>.90$

TABLE 6. SEX RATIOS AMONG OFFSPRING OF PRESUMABLY NORMAL PARENTS IN THE UTAH KINDRED

Parents	Sons	Daughters	Unk.	Total	Ratio M:F
13 Mothers	27	22	0	49	1.22
3 Fathers	3	6	0	9	0.50
16 Parents	30	28	0	58	1.07

It is clear in the last line of Table 4 that there is a sizeable deficiency of male progeny when all persons are considered (76 − 52 = 24). This deficiency of 24 males can be seen to be in the category of affected males, because the addition of 24 males in this category creates an almost perfect ratio for an autosomal dominant (31:37:32:37). It should be emphasized that this deficiency of affected males is the sum of separate deficiencies from *both affected fathers and affected mothers*.

Table 5 contains χ^2 tests of significance of the various categories of children of affected persons, assuming that the disorder is actually transmitted as a fully penetrant autosomal dominant. The deficiency of affected males is clearly seen to be the factor which has caused the significant disturbances in the various ratios. The heterogeneity becomes insignificant, for example, in the last line when the deficiency of 24 affected males is replaced. It is important to emphasize that the deficiency is of *affected males not affected females*. If a deficiency of affected females were observed, lack of penetrance in the asymptomatic carriers might be suspected. However, the fact that the deficiency has occurred in the sex with the more severe clinical disease suggests that the deficiency has resulted from a significantly higher death rate at some point in gestation among the affected males.

Table 6 records the sex ratios among the presumably unaffected members of the kindred. The usual slight excess of males over females is found in this part of the kindred. Thus it becomes very difficult to attribute the reversed sex ratio among affected persons to some tendency in the kindred not related to the kidney disease.

SUMMARY

Analysis of the Utah kindred with chronic kidney disease and nerve deafness shows it to be comparable in most respects to other pedigrees in the literature. It is, however, incomparable in extent and thoroughness of study. This kindred thus provides unique insights into the nature of hereditary chronic renal disease.

The kindred appears to be abnormal in several respects not emphasized by the authors, i. e.

1. There is a significant deficiency of males in the entire kindred, and

2. There is a deficiency in Generation IV of affected grandsons but not affected granddaughters from the affected grandmothers of Generation II. The deficiency has occurred, because *both the affected fathers and affected mothers of Generation III failed to produce the expected numbers of affected sons.*

I suggest the following hypothesis as more satisfactory than dominant partial sex-linkage for explaining the inheritance of the syndrome in this family.

1. It is transmitted as a sex-influenced, dominant trait, much more severe in males (to the point that 50% of affected males are never seen, presumably because of death *in utero*). Sex limitation is not complete, however, because some transmitting females show the defect.

2. The mutant gene is pleiotropic, incompletely penetrant (penetrating less frequently in females than males) and resides on an autosomal chromosome.

This complex hypothesis of autosomal transmission fits the observations better than does dominant partial sex-linkage. For example, the deficiency of affected sons from affected fathers is explained by either hypothesis. Only the autosomal hypothesis with the added modifications, however, explains also the discrepancy of sex ratio and the deficiency of affected grandsons from both affected mothers and affected fathers. The most appealing feature of the hypothesis, of course, is that it is consistent with the cytological observations regarding the X and Y chromosomes in humans.

REFERENCES

1. Alport, A. C. 1927. Hereditary familial hemorrhagic nephritis. *Brit. M. J.* 1: 504–506.
2. Eason, J., Smith, G. L. M., and Buchanan, G. 1924. Hereditary and familial nephritis. *Lancet* 2: 639–646.
3. Goldbloom, R. B., Fraser, F. C., Waugh, D., Aronovitch, M., and Wiglesworth, F. W. 1957. Hereditary renal disease associated with nerve deafness and ocular lesions. *Pediatrics* 20: 241–252.
4. Haldane, J. B. S. 1936. A search for incomplete sex-linkage in man. *Ann. Eugen.* 7: 28–57.
5. Hamburger, J., Crosnier, J., Lissac, J., and Naffah, J. 1956. Sur un syndrome familial de nephropathie avec surdite. *J. urol. méd.* 62: 113–124.
6. Morton, N. E. 1957. Further scoring types in sequential linkage tests with a critical review of autosomal and partial sex-linkage in man. *Am. J. Human Genet.* 9: 55–75.
7. Perkoff, G. T., Stephens, F. E., Dolowitz, D. A., and Tyler, F. H. 1951. A clinical study of hereditary interstitial pyelonephritis. *Arch. Int. Med.* 88: 191–200.
8. Perkoff, G. T., Nugent, C. A., Jr., Dolowitz, D. A., Stephens, F. E., Carnes, W. H., and Tyler, F. H. 1958. A follow-up study of hereditary chronic nephritis. *Arch. Int. Med.* 102: 733–746.
9. Reyersbach, G. C., and Butler, A. M. 1954. Congenital hereditary hematuria. *N. England J. M.* 251: 377–380.
10. Robin, E. D., Gardner, F. H., and Levine, S. A. 1957. Hereditary factors in chronic Bright's disease: a study of 2 affected kindreds. *Tr. Ass. Am. Physicians.* 70: 140–147.
11. Stephens, F. E., Perkoff, G. T., Dolowitz, D. A., and Tyler, F. H. 1951. Partially sex-linked dominant inheritance of interstitial pyelonephritis. *Am. J. Human Genet.* 3: 303–313.
12. Sturtz, G. S., and Burke, E. C. 1956. Hereditary hematuria, nephropathy and deafness. *N. England J. M.* 254: 1123–1126.

Discrimination of Genetic Entities in Muscular Dystrophy

C. S. CHUNG AND N. E. MORTON

Department of Medical Genetics, University of Wisconsin

OFTEN GENETIC ENTITIES differ in several attributes, none of which is an absolutely accurate diagnostic criterion. Thus autosomal and sex-linked forms of gargoylism have different frequencies of various symptoms (Herndon, 1954), and the recognition of heterozygotes in phenylketonuria may require determination of serum levels at intervals before and after administration of phenylalanine (Knox and Messinger, 1958). In such cases some measurements will be more critical than others, and it is natural to inquire what combination of measurements will best discriminate the genetic groups (Fisher, 1938; Bartlett, 1951). We here apply the method of discriminant functions to a large body of data on various types of muscular dystrophy, to clarify the inheritance of this disease and, in the process, to develop the applications of discriminant function analysis in human genetics.

METHODS

Classification problems are of two kinds, identification of group constellations and discrimination of groups. In the first case (Bartlett, 1951), k families are examined for heterogeneity by testing whether they fall into a small number less than k of phenotypic categories, which may be genetic entities. In the second case (Fisher, 1938), certain genetic categories are recognized, and the problem is to construct discriminant functions for classifying individuals or families into these groups.

Suppose that each individual is characterized by p variates, and let X_{hij} denote the h^{th} variate of the j^{th} individual in the i^{th} group ($h = 1, \cdots, p$; $i = 1, \cdots, k$; $j = 1, \cdots, n_i$). Also let $\sum_i n_i = N$ and $\{\sum n_i - \sum n_i^2/\sum n_i\}/(k-1) = c$. Associated with these quantities are two symmetric $p \times p$ matrices, W and B. W comprises the sums of squares and products within groups, which for the g^{th} and h^{th} variates is

$$w_{gh} = \sum_i \sum_j X_{gij} X_{hij} - \sum_i \frac{\sum_j X_{gij} \cdot \sum_j X_{hij}}{n_i},$$

with $N - k$ degrees of freedom. B is the sum of squares and products between groups, the corresponding element of which is

$$b_{gh} = \sum_i \frac{\sum_j X_{gij} \cdot \sum_j X_{hij}}{n_i} - \frac{\sum_{i,j} X_{gij} \cdot \sum_{i,j} X_{hij}}{\sum_i n_i},$$

Received June 15, 1959.

Genetics paper number 724. This work was supported by grants from the National Institutes of Health, Muscular Dystrophy Associations of America, and the Wisconsin Alumni Research Foundation.

with $k - 1$ degrees of freedom. Let w^{gh} be the element corresponding to w_{gh} in W^{-1}, the inverse of W. To test the null hypothesis that the mean values are the same in all k populations for these p variates, use

$$\chi^2_{p(k-1)} = (N - k) \sum_{g,h=1}^{p} w^{gh} b_{gh} .$$

To test whether q additional variates significantly increase the discriminatory power of the basic set of p variates, use

$$\chi^2_{q(k-1)} = (N - k) \sum_{g,h=1}^{p+q} w^{gh'} b_{gh'} - \chi^2_{p(k-1)} ,$$

where $w^{gh'}$ and $b_{gh'}$ are elements of matrices for the $p + q$ variates (Rao, 1952, chap. 7).

Discriminant functions are useful for investigation of differences among the k groups. The best function is linear if the variates are normally distributed and their variances and covariances are homogeneous among groups. The best function is quadratic if the variates are normal but not homoscedastic. With no assumption about normality or homoscedasticity, the coefficients of the linear function which maximizes the intraclass correlation may be denoted by a vector V_1 corresponding to the largest root λ_1 of the determinantal equation $| B - \lambda W | = 0$, the roots of which are the eigenvalues of the matrix BW^{-1}. The intraclass correlation corresponding to the r^{th} root λ_r is

$$\rho_r = \frac{\lambda_r(N - k) - (k - 1)}{\lambda_r(N - k) + (c - 1)(k - 1)} .$$

In large sample theory for $p > r$, a lower confidence limit for λ_r is given by $\lambda_r = e^{-\chi^2/m}(1 + \lambda_r^*) - 1$, where λ_r^* is the observed value, χ^2 is the critical value of chi-square for any specified confidence level and $p - r$ degrees of freedom, and $m = N - r - 1 - (p - r + k)/2$ (Bartlett, 1951).

The probability of resolving genetic entities is maximal for the largest root λ_1, which corresponds to the greatest intrafamily correlation. An intrafamily correlation significantly greater than ½ is suggestive of a mixture of genetic entities, but common environment, recording bias, or assortative mating could in principle give larger intrafamily correlations. Even a smaller correlation would suggest admixture if the groups were large kindreds of remote relatives or if the proportion of admixture were small, providing the distribution of the discriminant function were clearly bimodal. Such significant bimodality of individuals and family means is the only critical evidence for a mixture of phenotypic categories, which may be genetic entities. The converse is not true: unimodality and an intrafamily correlation less than ½ do not prove that there is not a mixture of genetic entities, only that they cannot be discriminated by the available data.

If genetic groups can be delimited accurately in certain cases, these may be made the basis for classifying other individuals and families. The appropriate method is to make pairwise comparisons of groups, so that the problem reduces to linear regres-

sion on preassignable group scores, and the above χ^2 test can be replaced by its variance ratio analogue. This procedure is given in detail by Rao (1952, chap. 8).

To exclude one group where there are only two alternatives, without assuming normality or homoscedasticity of the discriminant, calculate the mean score for the mixed group, say d. Then the expected frequency of group 1 in the mixture is $p_1 = E\{(d - u_2)/(u_1 - u_2)\}$, where u_1, u_2 are the means of the two unmixed groups. If m individuals are assigned their family means, and if $u_1 < u_2$, the $p_1 m$ individuals with the smallest scores should be classified as group 1, and the remainder as group 2. If ϵ_{12} is the proportion of group 1 among cases classified as group 2, and ϵ_{21} the proportion of group 2 among cases classified as group 1, then the total probability of error is $p_1\epsilon_{21} + p_2\epsilon_{12}$.

THE DATA

Although several studies have been done on the genetics of muscular dystrophy, the classification of clinical and genetic types is still in dispute. We have therefore thought it advisable, while analysing inheritance factors in our own material, to apply some of the same methods of analysis to comparable studies in the literature with the hope that a consistent genetic pattern could be discerned.

The usable data consist of eight sources: Sjovall (1936) from Sweden, Levison (1951) from Denmark, Stevenson (1953, 1954) from Northern Ireland, Walton (1955) from Northumberland and Durham, de Grouchy (1953) and Lamy and de Grouchy (1954) from France, Becker (1953) from Germany, our material from Wisconsin, and a few exceptionally thorough case reports from the medical literature (see references) which satisfied the requirements of careful neurological documentation and adequate family information. The extensive compilation of Julia Bell (1943) was found to be unsatisfactory in both respects but was helpful in suggesting hypotheses to test in our material.

Our Wisconsin data were collected in conjunction with a Muscular Dystrophy Clinic directed by Dr. H. A. Peters and sponsored by the Muscular Dystrophy Associations of America. Patients and their families came to the clinic at the University Hospitals for neurological examination and functional evaluation, supported by electromyography, muscle biopsy, and laboratory tests as indicated. Most patients were examined several times during a period of more than a year. On these visits detailed family histories were taken by the authors, and blood samples were obtained from patients and their relatives for serum aldolase and transaminase determinations.

These studies exclude neurogenic, myositic, ocular, distal, and myotonic myopathies and include the Duchenne, limb-girdle, and facioscapulohumeral types of Walton and Nattrass (1954). We have made no initial division into clinical groups, relying on the analysis to define entities within the category of muscular dystrophy. In view of evidence for dominant, recessive, and sex-linked types of dystrophy, a preliminary classification of pedigrees was made into five groups:

1. parent and child affected (possibly dominant)
2. not so; an affected girl in the pedigree (possibly recessive)
3. not so; maternal male relative affected (possibly sex-linked)

4. not so; only brothers affected (possibly autosomal or sex-linked)
5. ambiguous or doubtful

Group 5 comprises three pedigrees with both parent and child affected, but no other direct transmission in the kindred, which on the basis of a pilot analysis appeared to be phenotypically distinct from group 1 cases and similar to group 2. Of these, Sjovall family 99 is a consanguineous marriage of an affected woman which produced an affected son and daughter, Stevenson family 29 has an affected father and daughter, and Walton family L4 has an affected mother with two affected sons. The last two families were considered by their authors to be of the limb-girdle type with pelvifemoral onset and no facial weakness. We believe that these pedigrees represent either heterozygous expression of a recessive gene or marriages of individuals homozygous for a recessive gene with allelic heterozygotes. In a fourth pedigree, Becker family B33, the evidence for this hypothesis is stronger, since the affected progenitor married a woman who was both his first and second cousin, and had several affected children, males and females, of the same phenotype as group 2. These children had 14 progeny and 8 grandchildren, none with dystrophy. We have therefore classified this pedigree in group 2.

Before undertaking a formal genetic analysis, three questions must be answered:

1. Are the three presumptive genetic groups 1, 2, and 3 phenotypically distinguishable?
2. If so, is group 4 a separable mixture of genetic groups 2 and 3?
3. Can groups 1, 2, and 3 be subdivided into more homogeneous categories, which are genetic entities?

The material for analysis comprises 95, 214, 93, and 377 cases of groups 1, 2, 3, and 4, respectively, each with information on age of onset, localization of first symptoms, age and condition at death or last examination, pseudohypertrophy of calves, and other matters. Eighteen variables, $X_1, X_2, \cdots, X_{18}$ have been used in the analysis:

X_1 the pedigree group (1, 2, or 3)
X_2 sex; 0 for male, 1 for female
X_3 proband; 0 for secondary case, 1 for proband
X_4 inbreeding coefficient
X_5 age in years at death or last examination

onset

X_6 age at onset of first definite symptoms
X_7 facial onset; 0 for no facial onset, 1 for onset as weakness of any facial muscles (cannot smile, close eyelids, wrinkle brows, tapir mouth, etc.)
X_8 scapulohumeral onset as difficulty elevating arms, combing hair, etc. (0, 1)
X_9 pelvifemoral onset as waddling gait, stumbling, difficulty climbing stairs, etc. (0, 1)

condition at death or last examination

X_{10} dead (0, 1)
X_{11} facial weakness (0, 1)
X_{12} scapulohumeral weakness (0, 1)

X_{13} pelvifemoral weakness (0, 1)
X_{14} pseudohypertrophy or hypertrophy of calves, at any time during course of the disease (0, 1)
X_{15} contractures of Achilles tendon, invertors of foot, hamstrings, biceps brachii, or pronators of forearm (0, 1)
X_{16} confined to bed or chair (0, 1)
X_{17} obesity (0, 1)
X_{18} skeletal deformity indicated by pes cavus, scoliosis, severe lordosis, or fracture from slight trauma (0, 1)

Inclusion of products of X_5 with several of the other variables did not significantly increase discrimination, and therefore this part of the analysis will not be presented.

Phenotypic differences among groups 1, 2, and 3

For each pair of groups (1, 2), (2, 3), and (1, 3) the regression of X_1 on the other variates within the eight sources of data was computed. Retaining X_2 (sex), X_3 (proband), and X_{10} (dead), the variates which did not contribute significantly to the discrimination were deleted until a basic set of variates was isolated which accounted for all of the statistically significant discrimination. This discriminant was adjusted

TABLE 1. DISCRIMINATION BETWEEN GROUPS 1 AND 3

	Onset		Facial affection X_{11}	Pseudo-hypertrophy X_{14}	Score d	Number of Cases	
	Scapulo-humeral X_8	Pelvi-femoral X_9				Group 1	Group 3
Group 3 phenotype	0	1	0	1	1.39	—	67
	1	1	0	1	1.11	—	2
	0	1	1	1	.97	—	4
	0	1	0	0	.86	2	18
	1	1	0	0	.58	—	1
Group 1 phenotype	0	1	1	0	.45	5	1
	1	1	1	0	.17	4	—
	1	0	1	1	−.06	1	—
	1	0	0	0	−.17	14	—
	0	0	1	0	−.31	15	—
	1	0	1	0	−.59	54	—
Mean						95	93
Group 1	.768	.116	.832	.011			
Group 3	.032	1.000	.054	.785			

Presence of an attribute is denoted by 1, its absence by 0.

$$d = .1086 - .2820\,X_8 + .7560\,X_9 - .4156\,X_{11} + .5254\,X_{14}$$

	d.f.	s.s.	m.s.	F	P
X_2, X_3, X_{10}	3	54.501	18.167	189.24	<.001
X_8, X_9, X_{11}, X_{14} after fitting above	4	66.705	16.676	173.71	<.001
Residual variates	9	1.353	.150	1.56	.2–.3
Remainder within classes	164	15.802	.096		

to the standard of a living male proband by substituting $X_2 = 0$, $X_3 = 1$, $X_{10} = 0$ in place of the observed values. For each pair of groups the discriminatory value of the observations rests in a few variables, the residual data contributing no significant additional information.

Groups 1 and 3 are virtually discrete (table 1). Group 1 corresponds to the facioscapulohumeral clinical type, usually with facial affection and facial or scapulohumeral onset. Pelvifemoral onset is rare, and pseudohypertrophy extremely rare. Group 3 corresponds to the Duchenne clinical type, usually with pseudohypertrophy and always with pelvifemoral onset. Weakness of the lower face is a terminal symptom in a small proportion of cases. The two patients of group 1 who approach group 3 in phenotype were recorded as having pelvifemoral onset and no facial affection. They occur in typical facioscapulohumeral pedigrees with dominant inheritance (Levison 2-8 and 3-6). The only case of group 3 which approaches group 1 in phenotype is Becker B64, with facial involvement and no pseudohypertrophy. The other scored member of this pedigree is more typical, and the family mean lies outside the range of group 1. Therefore all pedigrees of groups 1 and 3 in this material, and nearly all cases, may be separated correctly by clinical criteria alone.

There is some overlap between groups 1 and 2 (table 2). Facial onset is never found in group 2, and facial affection is present in only a small proportion of cases. Pelvifemoral affection is invariable, and pelvifemoral onset the rule. Overlap between groups 1 and 2 is due to group 1 patients with pelvifemoral onset or no significant facial symptoms (17 per cent of group 1 individuals but only 5 per cent of cases classified by their family means) and to a few group 2 patients with scapulohumeral

TABLE 2. DISCRIMINATION BETWEEN GROUPS 1 AND 2

	Onset		Affection		Score d	Number of Cases	
	Facial X_7	Pelvi-femoral X_9	Facial X_{11}	Pelvi-femoral X_{12}		Group 1	Group 2
Group 2 phenotype	0	1	0	1	.56	2	174
	0	0	0	1	.25	5	14
	0	1	1	1	.16	9	14
Group 1 phenotype	0	0	0	0	.09	9	5
	0	0	1	1	−.16	26	6
	0	0	1	0	−.31	25	1
	1	0	1	1	−.32	14	—
	1	0	1	0	−.47	5	—
Mean						95	214
Group 1	.200	.116	.832	.589			
Group 2	0	.879	.098	.972			

$$d = .0915 - .1585\,X_7 + .3145\,X_9 - .4039\,X_{11} + .1545\,X_{13}$$

	d.f.	s.s.	m.s.	F	P
X_2, X_3, X_{10}	3	5.601	1.867	29.64	$<.001$
X_7, X_9, X_{11}, X_{13} after fitting above	4	27.633	6.908	109.65	$<.001$
Residual variates	10	.461	.046	.73	$>.9$
Remainder within classes	284	17.795	.063		

TABLE 3. DISCRIMINATION BETWEEN GROUPS 2 AND 3

	Inbreeding X_4	Pseudo-hypertrophy X_{14}	Contractures X_{15}	Score d	Number of Cases		
					Group 2	Group 3	Group 4
Group 3	1/32 or less	1	1	.23 or more	17	31	121
Phenotype	1/64 or less	1	0	.18 or more	46	42	92
Group 2	Yes or no	0	1	.12 or less	29	12	51
Phenotype	1/32 or more	1	0	.11 or less	6	—	2
	1/16 or more	1	1	.10 or less	0	—	1
	Yes or no	0	0	.01 or less	116	8	110
Mean					214	93	377
Group 2	.0067	.322	.215				
Group 3	0	.785	.462				
Group 4	.0018	.573	.459				

$$d = .0096 - 4.0356\,X_4 + .2306\,X_{14} + .1136\,X_{15}$$

	d.f.	s.s.	m.s.	F	P
X_2, X_3, X_{10}	3	20.837	6.946	80.77	$<.001$
X_4, X_{14}, X_{15} after fitting above	3	6.898	2.299	26.73	$<.001$
Residual variates	10	2.225	.223	2.59	.1–.2
Remainder within classes	283	24.405	.086		

onset or no pelvifemoral affection (6 per cent of group 2). Pedigrees and cases of groups 1 and 2 can usually, but not always, be separated correctly by clinical criteria alone.

Between groups 2 and 3 there is considerable overlap (table 3). Group 2 is characterized by a high frequency of consanguineous parentage and relatively low frequencies of pseudohypertrophy and contractures. In group 3 there is no increase of consanguineous parentage over the general population, the majority of patients have a history of pseudohypertrophy, and most of them develop contractures during the course of the disease. However, in group 2, 63/214 cases overlap group 3 by having pseudohypertrophy without close consanguinity, while 20/93 cases of group 3 give no history of pseudohypertrophy. Classified by their family means, 34 per cent of group 2 and 11 per cent of group 3 overlap the other group. Without genetic information, pedigrees and cases of groups 2 and 3 cannot always be separated correctly by clinical criteria alone.

In interpreting the above results, it is important to recognize that the significant clinical observations are to a degree subjective. Some patients are uncertain about the initial symptoms, and onset in the upper girdle or face may not be noticed until the patient has difficulty in walking. The presenting symptom is often not the initial symptom. Facial affection is notoriously difficult for the patient to remember and an observer to judge. Walton (1955) comments on the unreliability of family histories in group 1 and gives this anecdote from his own experience: "It was only at a later stage of the second interview that the author's suspicions of facial involvement in II.4 were aroused and detailed examination subsequently revealed that she was undoubtedly an abortive case. She had insisted previously that all her sibs were perfectly

well; however, when they, too, agreed to be examined, it was discovered that II.9 was another abortive case, and that II.7 was quite severely affected. In view of these findings it is certainly possible that either I.1 or I.2, both of whom died many years ago, were affected." The rather ill-defined facial weakness of some terminal group 3 cases, the inexpressive faces of many immobilized cases, and the occasional patients who cannot whistle, present different kinds and degrees of possible weakness about which observers may disagree.

Error is even more likely for mild contractures and pseudohypertrophy, which often go unremarked. The classical type of enlarged calves as seen in most group 3 cases is easy to recognize, but slight enlargement of the calves can be counterfeited by a stocky build or wasting of the vasti, while some terminal cases with reduced pseudohypertrophy could easily be missed. Reports by relatives on conformation of the calves are particularly unreliable. The primary importance attached to pseudohypertrophy by the discriminant between groups 2 and 3 must reflect a basic difference between group 3 and most of the members of group 2, but this cannot help but be obscured by errors in recognizing pseudohypertrophy. The existence of true pseudohypertrophy in some group 2 families was reported by Walton (1955), de Grouchy (1953), Kloepfer and Talley (1958), and others, and autosomal pseudohypertrophic muscular dystrophy has been demonstrated in mice (Michelson et al., 1955) and chickens (Asmundson and Julian, 1956). However, from several of our group 2 cases where pseudohypertrophy was debated and finally dismissed, we suspect that some of the published records of enlarged calves in group 2 are in error, while many of the terminal group 3 cases without a history of pseudohypertrophy may have manifested the condition earlier. With more refined clinical and laboratory criteria, groups 1, 2, and 3 are probably discrete.

Despite the crudeness of these data, the discriminant method has certain advantages over more impressionistic classifications. Given the observations, it is entirely objective, and permits genetic analysis even of studies which used different clinical categories. It focusses attention on variables with the greatest discriminatory value, showing that accuracy in their determination is important for correct group classification. The discriminant can be averaged over all patients in a pedigree, so that families may be classified by their means rather than, less reliably, by individual phenotypes. Using the discriminant we can separate group 4 cases into cases which are probably of genetic type 2 and those which are probably of type 3, with known frequencies of error. This is essential for a formal genetic analysis.

Division of group 4

The mean discriminant between groups 2 and 3 is .0814 for group 2, .2431 for group 3, and .1872 for group 4, so the proportion of type 2 cases in group 4 is approximately

$$\frac{.1872 - .2431}{.0814 - .2431} = .346$$

There are 377 group 4 scores, which should therefore be divided into about (.346)(377) = 130 cases of type 4(2) and 247 cases of type 4(3). When the family

means are ranked from smallest to largest, the 130th case has the value .1232, corresponding to a patient with contractures but without pseudohypertrophy or consanguinity. On the basis of their family means, 44 cases have the same value, 16 of which would be classified as 4(2) and 28 as 4(3) if the 130th case were the dividing point. To avoid this arbitrary choice, consideration was given to the following facts. The mean score for group 3 is probably smaller than for 4(3), since group 3 contains more dead members whose enlarged calves are often not recorded, and therefore the proportion of 4(2) may well be greater than .346. A family mean greater than .1232 can be attained only if at least one member has pseudohypertrophy, which is the rule for group 3 and the exception for group 2. Bearing these circumstances in mind, all group 4 cases with family means greater than .1232 were assigned to group 4(3), leaving 158 cases with family means .1232 or less assigned to 4(2), or $p_{4(2)} = .419$.

Let α_{23} be the proportion of genetic type 2 cases in group 4 that are misclassified as 4(3), and α_{32} be the proportion of genetic type 3 cases in group 4 that are misclassified as 4(2). From the overlap of groups 2 and 3, we estimate $\alpha_{23} = .34$, $\alpha_{32} = .11$. In the group classified as 4(2), the proportion actually of genetic type 3 is

$$\epsilon_{32} = \frac{p_{4(3)}\,\alpha_{32}}{p_{4(3)}\alpha_{32} + p_{4(2)}(1 - \alpha_{23})} = .19$$

and in the group classified as 4(3), the proportion actually of genetic type 2 is

$$\epsilon_{23} = \frac{p_{4(2)}\,\alpha_{23}}{p_{4(2)}\alpha_{23} + p_{4(3)}(1 - \alpha_{32})} = .22$$

If 2 and 4(2) are pooled, the proportion of cases actually of type 3 is

$$E_{32} = \frac{\epsilon_{32}\,n_{4(2)}}{n_{4(2)} + n_2} = .08$$

where $n_2 = 214$, $n_{4(2)} = 158$ are the numbers of cases classified as 2 and 4(2), respectively. Similarly, if 3 and 4(3) are pooled, the proportion of cases actually of type 2 is

$$E_{23} = \frac{\epsilon_{23}\,n_{4(3)}}{n_{4(3)} + n_3} = .15$$

where $n_3 = 93$, $n_{4(3)} = 219$ are the numbers of cases in groups 3 and 4(3). To be useful for genetic analysis, only scored pedigrees should be pooled. The total probability of error in classifying group 4 as 4(2) and 4(3) is $p_{4(2)}\epsilon_{32} + p_{4(3)}\epsilon_{23} = .20$. When 3 and 4(3) are pooled, and 2 and 4(2) are pooled, the total probability of error is $(.20)(377)/684 = .11$. This could undoubtedly be reduced by more precise observations.

In our Wisconsin material, when serum aldolase levels are taken into account, there is little, if any, phenotypic overlap between groups 2 and 3. However, autosomal pseudohypertrophic muscular dystrophy in the mouse and chicken shows the same high levels of aldolasemia that characterize group 3 in man, so that the aldolase test, although of considerable diagnostic and prognostic value, probably cannot provide

perfect discrimination between classes. A report on serum enzymes in clinical and genetic types of muscular dystrophy is in preparation.

SUBDIVISION OF GROUPS BY FISHER'S DISCRIMINANT

Levison (1951) and Becker (1953) suggested that isolated cases with scapulohumeral onset are often sporadic, but Walton (1955), Stevenson (1953), and others considered that all limb-girdle cases, whether isolated or familial, are due to autosomal recessive genes.

To put these speculations to test, we calculated Fisher's discriminant between isolated and familial probands within sources for group 2. If some isolated cases are sporadic, of different origin from familial cases, they should have a lower inbreeding coefficient and might differ in phenotype. Table 4 shows that, in fact, isolated probands are significantly less inbred than familial probands, but are not clinically distinguishable. This modifies a preliminary report (Chung and Morton, 1958) that isolated probands have significantly higher age of onset than familial probands, the difference being no longer significant when analysed within sources. Association between isolated cases and scapulohumeral onset is not significant. The inbreeding coefficients and segregation data indicate that about 41 per cent of group 2 cases in the population are sporadic, the remainder being due to autosomal recessive genes (Morton and Chung, 1959). Sporadic cases constitute an etiological type which, although not even suspected on clinical grounds, can be recognized by purely genetic methods. Since sporadic cases do not appear to be transmissible, they are not dominant mutants. Some of them may be due to rare expression of dystrophy in heterozygous carriers.

Fisher's discriminant may also be used to examine sex differences. When group 1 cases are compared within sources (table 5), the mean age of onset for males (16.8) is significantly greater than the mean age of onset (13.7) for females, suggesting that puberty may be a factor predisposing to the initial manifestation of the disease. Among group 2 familial cases there is no apparent clinical difference between the sexes.

TABLE 4. DISCRIMINATION BETWEEN ISOLATED AND FAMILIAL PROBANDS IN GROUP 2

	d.f.	s.s.	m.s.	F	P
X_2, X_{10}	2	8.578	4.289	27.67	$<.001$
X_4 after fitting above	1	1.374	1.374	8.86	.01–.001
Residual variates	12	1.865	.155	1.00	.3–.4
Remainder	123	19.013	.155		

X_2, X_4, X_{10} denote sex, inbreeding, and prior death, respectively.

TABLE 5. DISCRIMINATION BETWEEN SEXES IN GROUP 1

	d.f.	s.s.	m.s.	F	P
X_3, X_{10}	2	.236	.118	.52	.4–.5
X_6 after fitting above	1	1.323	1.323	5.83	.01–.02
Residual variates	11	3.115	.283	1.25	.2–.3
Remainder	74	16.798	.227		

X_3, X_6, X_{10} denote proband, age at onset, and prior death, respectively.

SUBDIVISION OF GROUPS BY BARTLETT'S DISCRIMINANT

Discrimination among group 1 pedigrees reveals a high degree of resemblance of related cases (table 6). The intraclass correlation is .747, with a lower 95% confidence limit of .673. The excess over ½ may well be due to multiple alleles or alternative loci determining this type of dystrophy, but the possibility that the apparent similarity of relatives is partly the result of systematic biases in reporting onset cannot be excluded. Figure 1 gives the distribution of scores (d). The four cases with the smallest scores have pelvifemoral onset and contractures, while the remaining cases with scores less than .4 include all patients with contractures and no scapulohumeral symptoms at onset. Three modes are suggested, but members of the same pedigree are often found in two or more of the apparent groups. These modes disappear when we consider the pedigree means, instead of being accentuated as would be expected if they reflected a mixture of genetic entities. We conclude that group 1 pedigrees cannot be separated into more homogeneous categories on the basis of present evidence.

The resemblance between affected relatives is even more marked for group 2 cases, with an intraclass correlation of .940. Low scores are associated with late onset, advanced age at the time of examination, a low frequency of pseudohypertrophy, and no contractures. There is no clear separation of juvenile and adult onset, scapulohumeral or pelvifemoral onset, pseudohypertrophic or not, or mild and severe cases, but rather a continuum of symptoms, as would be expected from multiple alleles or alternative loci with overlapping effects.

Similarity of relatives is also high for sex-linked cases (group 3), the intraclass correlation being .854. Low values of the score are associated with early onset, contractures, and confinement to bed. One case, Levison 23-1, has an unusually large score, which is due to onset at age 35 and slow course of the disease. However, his two brothers and nephew were more typical, with onset in childhood. Whatever the cause of the mild expression of dystrophy in this patient, he does not represent a distinct genetic type, nor do the other mild cases appear to be clearly separable from the severe forms. Furthermore, early onset and rapid progression are not always associated, one patient with onset at age 1 not being bedridden on examination at age 44 (de Grouchy 8.P), while another patient with onset at age 26 was bedridden at 30 (Walton D6.II.9). Thus the mild and severe types of sex-linked dystrophy, which Becker (1957) recognized, appear to exist only as extremes of a continuum. Sex-linked hemophilia shows the same marked resemblance of family members in AHF concentration and clotting time, perhaps for the same reason, that the disease may be caused by any of several alleles or pseudoalleles in different pedigrees. It is likely that there are a number of possible sites for mutation at each locus, with somewhat different effects.

TABLE 6. BARTLETT'S DISCRIMINANT AMONG FAMILIES WITHIN GROUPS 1, 2, AND 3

	Group 1	Group 2	Group 3
Eigenvalue, λ_1	4.683	32.743	12.504
Intraclass correlation, ρ_1	.747	.940	.854
Lower 95% confidence limit to ρ_1	.673	.926	.809

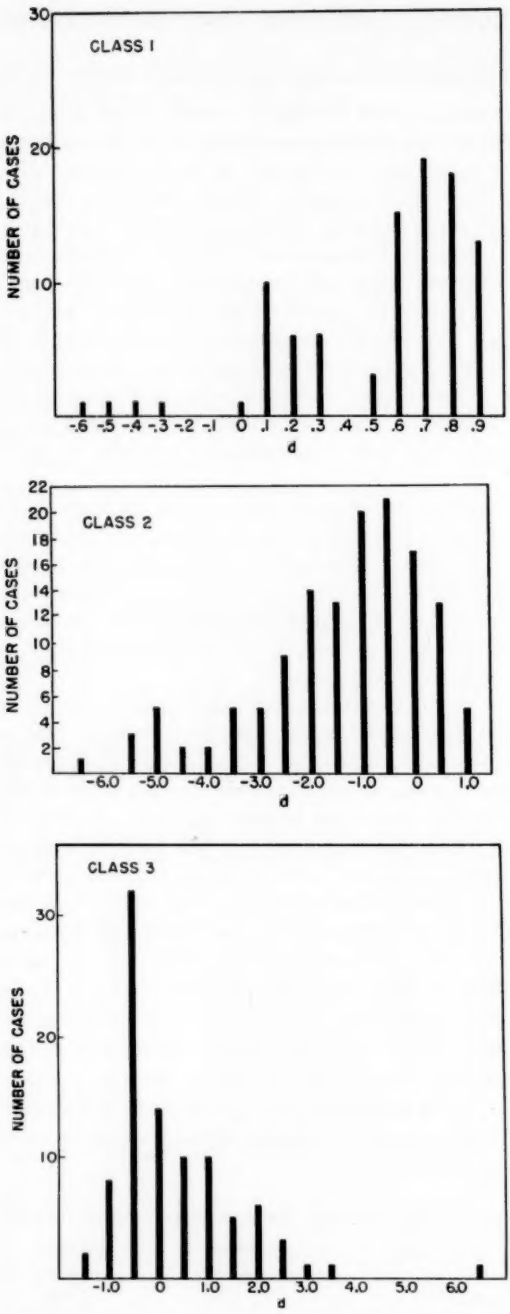

FIG. 1—Distribution of Bartlett's discriminant scores.

LINKAGE TESTS

Although groups 1, 2, and 3 do not appear to be subdivisible by clinical criteria, linkage analysis is one method that can resolve phenotypically indistinguishable genetic entities, as linkage to the Rh locus subdivides elliptocytosis (Morton, 1956). Linkage detection is most feasible for the dominant type (group 1), where large pedigrees may be studied. Linkage tests were not reported on the vast pedigree published by Tyler and Stephens (1950), but a few small pedigrees were tested by Walton (1955). Isolated cases of group 2 include a large proportion of sporadics, the genetic basis of which is unknown. Families with more than one case and/or parental consanguinity may be scored for linkage, but only a few informative families have been tested. Group 3 families are uninformative, since all the test factors are autosomal.

The small amount of available evidence is analysed in table 7 by probability ratio scores (Morton, 1955). Close linkage would be established by a score greater than 3, and reasonably excluded by a score less than -1. There is no evidence of close linkage to any test factor, but neither is there enough information to exclude moderate linkage.

DISCUSSION

Since the attempt to subdivide groups 1, 2, and 3 by clinical and genetic criteria has been unsuccessful, except for recognition of sporadic cases in group 2, we may profitably equate the three distinguishable classes to established clinical types. By combination of clinical and genetic criteria, patients can be assigned to these categories with little error. The following description of the major clinical types (excluding distal, ocular, myotonic, myositic, and neurogenic forms) is modified from Walton and Nattrass (1954).

The *Duchenne type* of muscular dystrophy is characterized by (a) transmission as a sex-linked recessive with a high mutation rate; (b) typical expression only in males;* (c) onset usually in the first 5 years of life, but occasionally as late as the third decade; (d) symmetrical involvement first of pelvic girdle muscles, spreading to the shoulder girdle usually within 3 to 5 years; (e) pseudohypertrophy of the calf muscles during the early course of the disease, sometimes disappearing in later stages; (f) steady and usually rapid progression without periods of apparent arrest, giving inability to walk usually within ten years of the onset but sometimes much later; (g) frequent rapid deterioration after periods of bedrest; (h) progressive deformity with muscular contractures and skeletal distortion and atrophy; (i) occasionally facial weakness in terminal stages; (j) death from inanition or respiratory infection usually in the second decade but sometimes not until middle life. This type comprises group 3, with definite evidence of sex-linkage, and 4(3), without such evidence but of the same phenotype as group 3.

---

* One recorded case in a pseudo-female with Turner's syndrome and male sex-chromatin (Walton, 1957). Homozygous females are a theoretical possibility, but must be fantastically rare; no well-documented case in a sex-linked pedigree has been reported. Apparently female heterozygotes may occasionally have atypical dystrophic symptoms (Kryschowa and Abowjan, 1934).

TABLE 7. LINKAGE TESTS

$\theta_1 = .10$

Test Locus	Source	Mating	Type	a	b	c	d	e	f	Score
					Group 1					
ABO	Walton, F1,II.4	$DdA_1 \times ddO$	(11)(11)	$2DA_1$	1DO	$-dA_1$	3dO	—	—	$z_1 = .2764$
	Walton, F2,I.1	DdB × ddO	(11)(11)	2DB	1DO	−dB	2dO	—	—	$z_1 = .0217$
	Walton, F2,II.5	DdB × ddA	(11)(11)	1DB	1DO	—	—	—	—	$z_1 = -.4437$
										$-.1456$
Rh	Walton, F2,II.2	$DR_1/dR_2 \times ddr$	(11)(11)	$1DR_1r$ $1DR_2r$ $1dR_1r$ $-dR_2r$ (1 n.c.o., 2 c.o.)						-1.1427
	Walton, F2,III.1	$DR_2/dr \times ddR_1r$	(11)(11)	$1DR_2/dR_1$ (1 n.c.o.)						.2553
	Walton, F2,II.5	$DR_1/dR_2 \times ddR_2r$		$1DR_1/dR_2$ $1DR_2/dR_2$ (1 n.c.o., 1 c.o.)						−.4437
										-1.3311
MN	Walton, F1,II.4	DdSs × ddss	(11)(11)	2DS	1Ds	−dS	3ds	—	—	$z_1 = .2764$
	Walton, F2,I.1	Multiple inconsistencies in published data, not explicable by illegitimacy								
	Walton, F2,II.5	DdMN × ddMN	(11)(121)	1DM	1DN	—	—			$z_1 = -.4473$
										$-.1673$
P	Walton, F1,II.4	DdPp × ddpp	(11)(11)	2DP	1Dp	1dP	2dp			$z_1 = -.6726$
Fy	Walton, F1,II.4	DdFf × ddff	(11)(11)	2DF	1Df	1dF	2df			$z_1 = -.6726$
	Walton, F2,I.1	DdFf × ddff	(11)(11)	2DF	1Df	1dF	1df			$z_1 = -.8874$
										-1.5600
Jk	Walton, F1,II.4	Ddab × ddab	(11)(121)	−Da	1Db	1da	2db			$z_1 = -.8874$
					Group 2 + 4(2)					
ABO	Stevenson D16	$DdA_1A_2 \times DdA_1B$	(31)(211)	$1DA_1$	$1dA_1$	$-DA_2B$	$1dA_2B$	$-DA_1B$	$1dA_1B$	$z_2(0, 1, 0, 1) = -.4437$, $e_2 = .1042$
										$z_2(1, 1, 0, 2) = -.3597$
	Stevenson D37	$DdBO \times DdA_2O$	(31)(1111)	$1DA_2B$	$1dA_2B$	$1DA_2$	$-dA_2$	1DB	1dB	$z_2(2, 2, 1, 0) = .0519$, $e_2 = -.0045$
										$z_2(2, 1, 1, 1) = -.4757$, $e_2 = -.0045$
	Stevenson D41	$DdA_1O \times DdO$	(31)(11)	$2DA_1$	$2dA_1$	1DO	−dO	—	—	$z_2 = .0519$, $e_2 = -.0045$
	Wisc. 62	$DdA_2O \times DdA_2O$	(31)(31)	$1DA_2$	2DO	$2dA_2$	−dO	—	—	$z_3 = .2268$, $e_3 = -.0091$
	Wisc. 316	$DdA_1B \times DdO$	(31)(11)	$-DA_1$	$2dA_1$	4DB	−dB	—	—	$z_2 = .6208$
										$-.2461$
Rh	Stevenson D16	$DdR_1r \times Ddrr$	(31)(11)	$-DR_1$	$1dR_1$	1Dr	2dr			

	Wisc. 316	$DdA_1B \times DdO$	(31)(11)	$-DA_1$	$2dA_1$	4DB	$-dB$	—	—	$z_2 = .6208$ / $-.2461$

Rh	Stevenson D16	$DdR_1r \times Ddrr$	(31)(11)	$-DR_1$	$1dR_1$	1Dr	2dr			$z_2 = -.5479$
	Stevenson D22	$DdR_1r \times Ddrr$	(31)(11)	$-DR_1$	$2dR_1$	1Dr	$-dr$			$z_2 = .3153$
	Stevenson D37	$DdR_0r \times Ddrr$	(31)(11)	$1DR_0$	$-dR_0$	2Dr	2dr			$z_2 = .0519$
	Stevenson D41	$DdR_1r \times DdR_2r$	(31)(1111)	$3DR_1r$	$1dR_1r$	$-Drr$	1drr	—	—	$z_2(3, 1, 0, 1) = -.3597$
										$z_2(3, 2, 0, 0) = -.1678$
	Wisc. 58	$DdR_1R_2 \times DdR_1R_1$	(31)(11)	$-DR_1R_1$	$-dR_1R_1$	$-DR_1R_2$	$2dR_1R_2$			$z_2 = .2148$
	Wisc. 62	$DdR_1r \times DdR_1r$	(31)(121)	$-Drr$	$-drr$	$2DR_1R_1$	$1dR_1R_1$	$1DR_1r$	$1dR_1r$	$z_3(1, 3, 1) = -.4741$
	Wisc. 316	$DdR_1R_2 \times DdR_1R_2$	(31)(121)	$-DR_2R_2$	$-dR_2R_2$	$3DR_1R_1$	$2dR_1R_1$	$1DR_1R_2$	$-dR_1R_2$	$z_3(1, 3, 2) = -.7915$
										-1.7590
MN	Stevenson D16	DdMN × DdN	(31)(11)	1DMN	3dMN	$-DN$	$-dN$	—	—	$z_2 = .3311$
	Stevenson D22	DdMN × DdM	(31)(11)	1DMN	$-dMN$	$-DM$	2dM			$z_2 = .3153$
	Stevenson D37	DdMN × DdM	(31)(11)	$-DMN$	1dMN	3DM	1dM			$z_2 = -.3597$
	Stevenson D41	DdMN × DdM	(31)(11)	$-DMN$	1dMN	3DM	1dM			$z_2 = -.3597$
	Wisc. 58	DdMN × DdM	(31)(11)	$-DMN$	1dMN	$-DM$	1dM			$z_2 = -.4437$
										$-.5167$
Fy	Wisc. 58	DdFf × Ddff	(31)(11)	$-DF$	1dF	$-Df$	1df	—	—	$z_2 = -.4437$, $e_2 = .1042$ / $-.3395$
K	Wisc. 58	DdKk × Ddkk	(31)(11)	$-DK$	$-dK$	$-Dk$	2dk	—	—	$z_2 = .2148$, $e_2 = .1042$ / $.3190$
PTC	Stevenson D16	DdTt × Ddtt	(31)(11)	$-DT$	2dT	1Dt	1dt	—	—	$z_2 = -.3597$, $e_2 = .0344$
	Stevenson D22	DdTt × Ddtt	(31)(11)	$-DT$	2dT	1Dt	$-dt$			$z_2 = .3153$, $e_2 = .0134$
	Stevenson D41	DdTt × Ddtt	(31)(11)	2DT	1dT	1Dt	1dt			$z_2 = -.4757$, $e_2 = -.0045$ / $-.4768$
Le	Stevenson D22	DdLl × Ddll	(31)(11)	$-DL$	2dL	1Dl	$-dl$			$z_2 = .3153$, $e_2 = .0134$
	Stevenson D37	DdLl × Ddll	(31)(11)	$-DL$	$-dl$	3Dl	2dl			$z_2 = -.1678$, / $.1564$ $e_2 = -.0045$
P	Stevenson D37	DdPp × Ddpp	(31)(11)	1DP	2dP	2Dp	$-dp$			$z_2 = .2832$, $e_2 = -.0045$ / $.2787$

The characteristic features of the *facioscapulohumeral form* are: (a) transmission as an autosomal dominant gene; (b) expression in either male or female; (c) onset usually in adolescence, but may be at any time from early childhood until middle age; (d) the occurrence of abortive cases; (e) initial or early involvement, sometimes asymmetrical, of facial and scapulohumeral muscles, usually with spread within twenty to thirty years to the pelvic girdle; (f) muscular pseudohypertrophy is extremely rare; (g) slow insidious progression with periods of long arrest of the disease; (h) severe disability does not usually occur and muscular contractures and skeletal changes are generally slight; (i) most patients survive and remain active to a normal age. This type comprises group 1, with evidence of dominant inheritance.

The *limb-girdle form* shows (a) transmission by an autosomal recessive gene in about 59 per cent of cases, sporadic occurrence of unknown etiology in the remainder; (b) expression in either male or female; (c) onset usually in childhood but sometimes in maturity or middle age; (d) primary involvement of either the pelvic or, less frequently, the shoulder girdle muscles, often with asymmetry of wasting when the upper limbs are first involved; (e) spread from the lower to the upper limbs and vice versa, in most cases within twenty years; (f) abortive cases are uncommon; (g) pseudohypertrophy of calves is uncommon, but may be counterfeited by a stocky build or wasting of the vasti; (h) variable rate of progression of the disease, fairly rapid in some, slow in others, sometimes with periods of clinical arrest; (i) severe disablement with inability to walk is usually seen within twenty to thirty years of the onset; (j) muscular contractures are uncommon save in the late stages and skeletal changes are slight; (k) sometimes facial weakness in late stages; (l) death frequently occurs in middle age when the disease begins in the second decade, but deaths in adolescence and old age are not unusual. This type comprises groups 2, with an affected girl; 4(2), without an affected girl but of the same phenotype as group 2; and 5, with an affected parent, but of the same phenotype as group 2 and with no conclusive evidence of dominant inheritance.

The different prognoses associated with these three types of dystrophy are illustrated by the onset distribution in figure 2 and especially by the survival curve in figure 3. The probability of survival to age x is $\sum_{i=x}^{\infty} n_i/N$, where N is the total number of dead patients and n is the number who died at age i. The death rate of facioscapulohumeral cases is not much different from the general population. The other two types have high mortality before age 20 when onset is in childhood, but a decreased death rate in adults. This may be related to the effect of the disease during puberty in causing contractures and impaired vital capacity. If the patient survives these hazards and completes his growth before disablement, the chance of his survival for the next 20 years or more is fairly good. This suggests the importance of physical therapy and other supportive measures, and stresses the need for occupational therapy. The classical description of the Duchenne type as "seldom compatible with adult life" (Tyler and Wintrobe, 1950) neglects the substantial proportion of cases with late onset and/or slow progression, who survive into middle age.

Lest there be any misunderstanding, we wish to emphasize that these categories are primarily genetic, since we consider the mode of inheritance to be the most

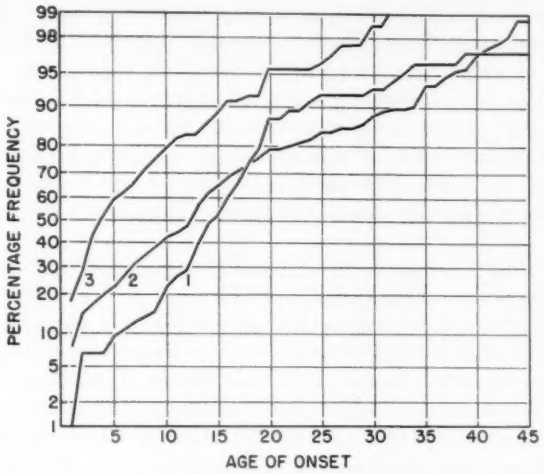

FIG. 2—Cumulative distribution of age of onset among affected persons.

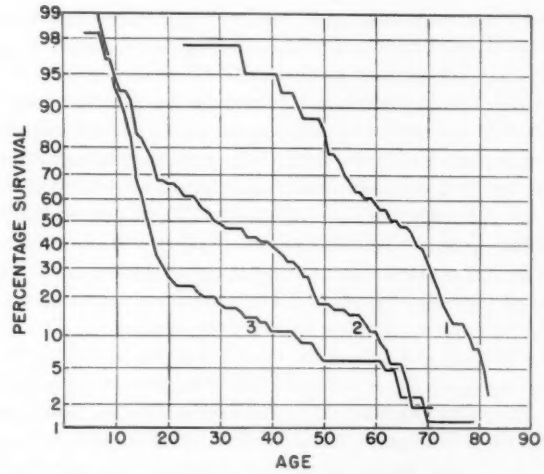

FIG. 3—Probability of survival until age X among dead cases.

characteristic feature of each type, foreshadowing fundamental biochemical differences to be discovered. If the inheritance is ambiguous, as in group 4, the probable genetic category is inferred from phenotypic evidence, guided by the discriminant score. On this basis, a patient with scapulohumeral onset, facial weakness, and unaffected parents would be classified as limb-girdle, as would a case of pseudohypertrophic dystrophy in a girl or in a child of consanguineous parents. It may be that there are subtle but definite differences between Duchenne and limb-girdle cases

with apparent pseudohypertrophy, as Becker (1953, 1957) and Stevenson (1953) believe, or that they are clinically indistinguishable, as Walton (1955) and de Grouchy (1953) claim. In either case, we prefer to reserve the appellation "Duchenne" for cases whose inheritance is probably sex-linked, and to avoid the term where there is a presumption of another mode of inheritance. Similarly, we use the designation "facioscapulohumeral" only for dominant cases, excluding limb-girdle cases with scapulohumeral onset and subsequent facial affection. This also excludes the few families in group 5, which are clearly limb-girdle even though they come from affected parents.

Stevenson (1953) has discussed the classification proposed by Bell (1943), which cuts across genetic and clinical types, and the attempt of Tyler and Wintrobe (1950) which omits the limb-girdle category. Levison (1951) and de Grouchy (1953) emphasized the difficulty of distinguishing between the limb-girdle and Duchenne types on clinical grounds, while Sjovall (1936) and Becker (1953) made a primary distinction between pelvic and shoulder girdle varieties. According to our studies and to those of Stevenson (1953) and Walton (1955), this subdivision of the limb-girdle type is not justified on present evidence.

Stevenson (1953, 1955) classified 33 families in Northern Ireland as Duchenne and 27 as limb-girdle. In his material he did not observe the dominant, facioscapulohumeral type. Of his Duchenne families, four are sex-linked and all are classified as Duchenne by the discriminant score. Of his limb-girdle families, 13 contain affected girls and all are classified as limb-girdle by the discriminant function, including one group 5 patient with an affected father. The only disagreement between his classification and ours is for family D17, which he classifies as Duchenne and the discriminant score identifies as limb-girdle. There is no pseudohypertrophy, contractures are not mentioned, and the proband's first cousin once removed, the child of his grandfather's brother, apparently had the same disease. These observations would seem to rule out sex-linkage and the Duchenne form.

Walton and Nattrass (1954) first proposed the trichotomous classification of the common types of muscular dystrophy, which we have found agrees best with the genetic evidence. Of the 18 families classified by Walton (1955) as the limb-girdle type, 7 have an affected girl and all are classified as limb-girdle by the discriminant score, including one group 5 case with an affected mother. However, of the 34 families considered by him to be Duchenne, 7 are clearly sex-linked, 16 others are classified as Duchenne by the discriminant, but 2 have an affected girl and 9 others are classified as limb-girdle by the discriminant. One of the affected girls had "very marked pseudohypertrophy of the calves" but no contractures, while the other had neither contractures nor pseudohypertrophy. None of the patients in the other families had pseudohypertrophy at the time of examination; there was no statement about previously enlarged calves. Evidently these cases were classified on the basis of onset in childhood and relatively rapid progression, but this is not a reliable discriminant between groups 2 and 3, and therefore not between Duchenne and limb-girdle cases. We believe that the discriminant score provides the most reliable clinical basis for classification of families in which the mode of inheritance is not known.

TABLE 8. CARDINAL SYMPTOMS OF THE MAJOR TYPES OF MUSCULAR DYSTROPHY

	Facioscapulohumeral Group 1	Limb Girdle Group 2		Duchenne Group 3
		Isolated	Familial	
Number of cases	95	79	135	93
Median age of onset	15	13	11	4
Mean inbreeding coefficient	0	.0012	.0080	0
Frequency of:				
Facial onset	.200	0	0	0
Scapulohumeral onset	.768	.215	.126	.032
Pelvifemoral onset	.116	.835	.904	1.000
Pseudohypertrophy of calves	.011	.316	.326	.785
Contractures	.095	.177	.237	.462
Facial affection	.832	.101	.096	.054
Scapulohumeral affection	1.000	.949	.919	.828
Pelvifemoral affection	.589	.949	.978	1.000
Confined to bed or chair	.074	.253	.341	.462

SUMMARY

Approximately 800 cases of muscular dystrophy from Wisconsin and from comparable studies in the literature were classified on pedigree information into three presumptive genetic groups and a fourth mixed group. Eighteen variables associated with mode of onset and clinical condition were used for discrimination among the three primary groups and to classify the mixed group. In one category, isolated probands have significantly less parental consanguinity than familial probands, but are not clinically distinguishable. The data indicate that about 59 per cent of the cases in this group are attributable to autosomal recessive genes and the remainder are sporadics of uncertain origin. Since the three primary groups could not be subdivided into more homogeneous genetic classes by the discriminant methods of Bartlett or by linkage analysis, they may be equated to established clinical types: dominant facioscapulohumeral, recessive and sporadic limb-girdle, and sex-linked Duchenne. Pedigree data and discriminant scores in conjunction provide the most accurate basis for classification of cases, with priority being given to the pedigree data.

ACKNOWLEDGMENTS

The authors wish to acknowledge the help given them by Dr. H. A. Peters and the staff of the Muscular Dystrophy Clinic in studying these families, by the Wisconsin State Medical Society and the Bureau for Handicapped Children in finding cases, and by the patients and their famiies whose cooperation made this study possible. We are grateful to Dr. Margery P. Gray for blood typing of the families used in the linkage analysis.

REFERENCES

ARBUSE, D. I. AND D. SLOANE. 1937. Muscular dystrophy; Report of a family. *N. York State J. M.* 37: 1111–1118.

ASMUNDSON, V. S. AND L. M. JULIAN. 1956. Inherited muscle abnormality in the domestic fowl. *J. Hered.* 47: 248–252.

BARTLETT, M. S. 1951. The goodness of fit of a single hypothetical discriminant function in the case of several groups. *Ann. Eugen.* 16: 199–214.

BATTEN, F. E. 1903. Two cases of muscular atrophy of uncertain origin. *Brain* 26: 151.

BECKER, P. E. 1953. *Dystrophia musculorum progressiva. Eine genetische und klinische Untersuchung der Muskeldystrophien.* Sammlung Psychiatrischer und Neurologischer Einzeldarstellungen. Stuttgart.

BECKER, P. E. 1957. Neue Ergebnisse der Genetik der Muskeldystrophien. *Act. Genet.* 7: 303–310.

BELL, JULIA. 1943. *The treasury of human inheritance.* Vol. IV. Nervous diseases and muscular dystrophies. IV. On pseudohypertrophic and allied types of progressive muscular dystrophy. Cambridge: University Press.

BREGMAN, L. E. 1899. Ein casuistischer Beitrag zur progressiven Muskelatrophie. *Deut. Zschr. Nervenh.* 14: 254–287.

CHUNG, C. S. AND N. E. MORTON. 1958. Discrimination of genetic entities in muscular dystrophy. *Proc. X Internat. Cong. Genet.* II: 51.

DE GROUCHY, J. 1953. *Contribution a l'Étude de l'Hérédité de la Myopathie.* Paris.

ERB, W. 1884. Ueber die juvenile Form der progressiven Muskelatrophie. *Deut. Arch. klin. Med.* 34: 467–519.

ESHNER,A. A. 1902. Two cases of progressive muscular dystrophy in brother and sister. *Philadelphia M. J.* 9: 1023–1025.

FISHER, R. A. 1938. The statistical utilization of multiple measurements. *Ann. Eugen.* 8: 376–386.

GROTHAUS, E. 1935. Untersuchungen über den Erbgang der Dystrophia musculorum progressiva (Erb). *Arch. klin. Chir.* 181: 344–362.

HERNDON, C. N. 1954. Genetics of the lipidoses. *Res. pub. Asso. Res. Nerv. Ment. Dis.* 33: 239–258.

JENDRASSIK, E. 1902. Beitrag zur Kenntniss der hereditären Krankheiten. *Deut. Zschr. Nerven.* 22: 444–499.

KLOEPFER, H. W. AND C. TALLEY. 1958. Autosomal recessive inheritance of Duchenne-type muscular dystrophy. *Ann. Hum. Genet.* 22: 138–143.

KNOX, W. E. AND E. C. MESSINGER. 1958. The detection in the heterozygotes of the metabolic effect of the recessive gene for phenylketonuria. *Am. J. Human Genet.* 10: 53–60.

KOLLARITS, J. 1906. Beiträge zur Kenntnis der vererbten Nervenkrankheiten. *Deut. Zschr. Nervenh.* 30: 293–363.

KOSTAKOW, S. AND K. BODARWÉ. 1938. Gibt es einen einheitlichen Erbgang bei der progressiven Muskeldystrophie? *Deut. Arch. klin. Med.* 181: 611–634.

KRYSCHOWA, N. AND W. ABOWJAN. 1934. Zur Frage der Heredität der Pseudohypertrophie Duchenne. *Zschr. ges. Neur. Psychiat.* 150: 421–426.

LAMY, M. AND J. DE GROUCHY. 1954. L'Hérédité de la myopathie (formes basses). *J. Génét. Hum.* 3: 219–261.

LEVISON, H. 1951. Dystrophia musculorum progressiva. Clinical and diagnostic criteria; inheritance. *Opera ex Domo Biologicae Hereditariae Humanae Universitatis Hafniensis.* 26.

MICHELSON, A. M., E. S. RUSSELL, AND P. J. HARMAN. 1955. Dystrophia muscularis: A hereditary primary myopathy in the house mouse. *Proc. Nat. Acad. Sc.* 41: 1079–1084.

MILLS, H. P., T. H. HAINES, AND M. A. SESSIONS. 1924. Familial muscular dystrophy. *Arch. Neur. Psychiat.* 11: 43–53.

MORTON, N. E. 1955. Sequential tests for the detection of linkage. *Am. J. Human Genet.* 7: 277–318.

MORTON, N. E. 1956. The detection and estimation of linkage between the genes for elliptocytosis and Rh blood type. *Am. J. Human Genet.* 8: 80–96.

MORTON, N. E. 1957. Further scoring types in sequential linkage tests, with a critical review of autosomal and partial sex linkage in man. *Am. J. Human Genet.* 9: 55–75.

MORTON, N. E. AND C. S. CHUNG. 1959. Formal genetics of muscular dystrophy. *Am. J. Human Genet.* 11: 360–379.

ORANSKY, W. 1927. Über einen hereditären Typus progressiver Muskeldystrophie. *Deut. Zschr. Nervenh.* 99: 147–155.

ORMEROD, J. A. 1903. Case of myopathy; Erb's juvenile type. *Brain* 26: 144–146.

PEARSON, K. 1933. Two pedigress of muscular dystrophy. *Ann. Eugen.* 5: 179–191.

RACE, R. R. 1955. On the inheritance of muscular dystrophy. A note on the blood groups. *Ann. Human Genet.* 20: 13–16.

RAO, C. R. 1952. *Advanced statistical methods in biometric research.* New York: Wiley and Sons.

RIESE, W. 1922. Über familiäre verebbare Dystrophien der Gesichtsmuskulatur. *Deut. Zschr. Nervenh.* 75: 214–223.

SJÖVALL, B. 1936. *Dystrophia musculorum progressiva. Eine erblichkeitsmedizinische und klinische studie.* Copenhagen: Levin and Munksgaard.

STEVENSON, A. C. 1953. Muscular dystrophy in Northern Ireland. I. An account of the condition in fifty-one families. *Ann. Eugen.* 18: 50–93.

STEVENSON, A. C. 1955. Muscular dystrophy in Northern Ireland. II. An account of nine additional families. *Ann. Human Genet.* 19: 159–164.

STEVENSON, A. C., E. A. CHEESEMAN, AND M. C. HUTH. 1955. Muscular dystrophy in Northern Ireland. III. Linkage data with particular reference to autosomal limb girdle muscular dystrophy. *Ann. Human Genet.* 19: 165–173.

STONE, S. 1940. Treatment of muscular dystrophies and allied conditions. *J. Am. M. Ass.* 114: 2187–2191.

TYLER, F. H. AND M. M. WINTROBE. 1950. Studies in disorder of muscle: the problem of progressive muscular dystrophy. *Ann. Int. M.* 32: 72–79.

WALTON, J. N. AND F. J. NATTRASS. 1954. On the classification, natural history and treatment of the myopathies. *Brain* 77: 169–231.

WALTON, J. N. 1955. On the inheritance of muscular dystrophy. *Ann. Human Genet.* 20: 1–38.

WALTON, J. N. 1956. The inheritance of muscular dystrophy: further observations. *Ann. Human Genet.* 21: 40–58.

WALTON, J. N. 1957. The inheritance of muscular dystrophy. *Act. Genet.* 7: 318–320.

WINOCOUROFF, J. 1907. Zwei Fälle von Dystrophia muscularis progressiva familiaris. *Arch. Kinderh.* 46: 109–112.

Formal Genetics of Muscular Dystrophy

N. E. MORTON AND C. S. CHUNG

Department of Medical Genetics, University of Wisconsin

SINCE ITS ORIGIN, human genetics has been subject to several sources of error, stemming both from the nature of the material and from the methods of analysis that have been applied. Two generations of geneticists have removed some of the technical limitations, but the full potentialities of genetic analysis in man have certainly not yet been realized.

These considerations led to the development of a rigorous theory of segregation analysis in man (Morton, 1958, 1959) and to the search for a trait that would test these methods to the uttermost. We required a large body of data on each of the simple modes of inheritance, made up in part of representative conventional studies, in part of the more complete data which precise analytical methods demand.

Muscular dystrophy meets these conditions admirably. Dominant, recessive, and sex-linked types are known, and a sporadic type was revealed by our investigations. A number of excellent studies have been published which contain sufficient information for genetic analysis (Sjovall, 1936; Stephens and Tyler, 1951; Tyler and Stephens, 1950; Levison, 1951; Stevenson, 1953, 1954; Walton, 1955; de Grouchy, 1953; Lamy and de Grouchy, 1954; and Becker, 1953). A small group of sibships with complete or truncate selection (Morton, 1958, 1959) and unusually good clinical and genetic data, compiled from the literature by Chung and Morton (1959), have been included as an additional sample. Casual reports from the literature, like those compiled by Julia Bell (1943), have heuristic value but, in the absence of adequate data on the number, age, and condition of relatives, mode of ascertainment, consanguinity, and other particulars, are not suitable for this kind of genetic analysis. Through the cooperation of Dr. H. A. Peters, director of the Muscular Dystrophy Clinic of the University Hospitals, we were able to collect genetic information that had never before been reported, including data on the numbers of ascertainments of probands and chemical studies on the various types of dystrophy (Chung, Morton, and Peters, in preparation). Analysis of the genetic categories by discriminant functions indicated that they could be recognized with relatively little error (Chung and Morton, 1959).

During the course of this work there was an exciting interplay between collection of data which required more refined methods of genetic analysis than were then known, and development of techniques that required data that were not then available but are included in our Wisconsin material. The result has been a stringent test of the potentialities of modern methods of analysis in human genetics.

Received January 10, 1959.

Genetics paper number 622. This work was supported by grants from the National Institutes of Health, Muscular Dystrophy Associations of America, and the Wisconsin Alumni Research Foundation.

Segregation analysis of class 1 (facioscapulohumeral)

In a previous paper Chung and Morton (1959) defined four genetic classes which will be used here.

Class 1 consists of cases with an affected parent, excluding four families where there is a strong presumption of recessive inheritance. Most of these cases are in large kindreds, with obviously dominant inheritance. Let there be r affected, a of them probands, in a sibship of size s. The probability of only one affected, if there is at least one affected, is

$$P(r = 1 \mid r > 0) = \frac{sp\pi[x + (1 - x)q^{s-1}]}{xsp\pi + (1 - x)[1 - (1 - p\pi)^s]} \quad \text{(problem 2)}$$

where p is the expected proportion affected, $q = 1 - p$, x is the proportion of cases that are sporadic, and π is the ascertainment probability, which is unity if ascertainment was through a parent or other relative outside the sibship (Morton 1958, 1959). To facilitate reference to this distribution, its solution by maximum likelihood scores will be called problem 2; other problems will be defined and numbered presently.

For simplicity we assume that p and π are constant among sources and that $x = 0$, sibships without an affected parent having been excluded. The value of $\pi < 1$ must be estimated from the data by methods to be considered, but the value of p is determined by genetic hypothesis. For a dominant gene without appreciable viability effects before the age of diagnosis, $p = y/2$, where y is the average penetrance in the sample, or the probability of being affected if susceptible. If all carriers who live long enough are recognizably affected, y lies between $\int f(z)G(z)\,dz$ and $\int f_1(z)G(z)\,dz$, where $f(z)$ is the frequency of age z at death or last examination among affected individuals and their normal sibs, $f_1(z)$ is this frequency with the first proband in each family excluded, and $G(z)$ is the cumulative frequency of onset at age z among affected cases (Morton, 1958). From the ages given in class 1 sibships, the calculated value of y is .908 for affected individuals who are not probands, .895 for all affected individuals, and .707 for their normal sibs, or approximately .804 for all sibs, giving $p = .402$.

Using the same symbols, the distribution of affected individuals in sibships with more than one affected is

$$P_{(r \mid r>1)} = \frac{\binom{s}{r} p^r q^{s-r}[1 - (1 - \pi)^r]}{1 - (1 - p\pi)^s - \pi spq^{s-1}} \quad \text{(problem 3)}$$

The distribution of probands among affected individuals is

$$P_{(a)} = \frac{\binom{r}{a} \pi^a (1 - \pi)^{r-a}}{1 - (1 - \pi)^r} \quad \text{(problem 9)}$$

while the distribution of $t > 0$ ascertainments among r affected is

$$P_{(t)} = \frac{[-r \ln(1 - \pi)]^t (1 - \pi)^r}{t![1 - (1 - \pi)^r]}. \quad \text{(problem 6)}$$

TABLE 1. SEGREGATION ANALYSIS OF CLASS 1

Source	Problem No.	p	x	π	U_p	U_π	K_{pp}	$K_{\pi\pi}$	$K_{p\pi}$
a > 0 probands	2	.402	0	.247	4.645	−0.579	98.57	1.61	−12.58
a = 0	2	.402	0	1	3.025	—	214.30	—	—
a > 0	3	.402	—	.247	−25.355	2.990	729.06	4.90	−57.11
a = 0	3	.402	—	1	70.294	—	1860.10	—	—
a > 0	9	—	—	.247	—	−20.345	—	99.18	—
Our data, a > 0	6	—	—	.247	—	−1.391	—	6.80	—
Walton, a > 0	6	—	—	.247	—	16.152	—	64.54	—
Total					52.609	−3.173	2902.03	177.03	−69.69
Sex of affected	10	½	—	—	48.000	—	1504.00	—	—
Sex of normal	10	½	—	—	−16.000	—	1896.00	—	—

	d.f.	χ^2	P
Deviation from p = .402	1	0.92	.5–.3
p among sources	3	2.85	.5–.3
p among samples	6	9.58	.2–.1
Sex of affected, p among samples	7	11.04	.2–.1

Becker (1953) distinguishes sibships with probands but not the probands themselves, hence neither problem 6 nor 9 is applicable. Walton (1955) reports ascertainments but not probands, while the remaining studies, with the exception of our Wisconsin material, indicate probands but not ascertainments. For our Wisconsin data problem 6 simplifies to the case r = 1, since we have recorded the number of ascertainments per proband. A preliminary estimate of π, made before the Wisconsin data were tabulated, was .247. The analysis by maximum likelihood scores, using p = .402, π = .247, is set out in table 1.

Assuming p = .402, an improved estimate of π is $.247 + U_\pi/K_{\pi\pi} = .229$, where U_π is the total score for π and $K_{\pi\pi}$ is its variance. At this value the score for p by Taylor's theorem is $U_{p|\pi} = .247 + (.247 - .229)K_{p\pi} = 52.609 + (.018)(-69.69) = 51.355$, the correction to U_p being negligible. The information matrix for p and π is

$$\begin{bmatrix} K_{pp} & K_{p\pi} \\ K_{p\pi} & K_{\pi\pi} \end{bmatrix},$$

which inverts to the covariance matrix

$$\begin{bmatrix} 3.4788 & 1.3695 \\ 1.3695 & 57.0267 \end{bmatrix} 10^{-4}.$$

To test the hypothesis that p = .402 when π is estimated from the sample, we use $\chi_1^2 = (51.355)^2(3.4788)10^{-4} = 0.92$. Clearly the data satisfy this hypothesis.

The largest estimate of p is given by the pedigree of Tyler and Stephens (1950), with $U_p = 84.253$, $K_{pp} = 1135.63$. These authors eliminated "families with many

children under twelve," giving p = ½ as the expected and observed frequency of affected. There is no significant heterogeneity among samples, nor do the other samples differ significantly from p = .402.

The distribution of r females among s affected individuals is $\binom{s}{r}p^r q^{s-r}$, where p = ½ on the null hypothesis of an equal sex ratio (problem 10). The data are clearly consistent with this, and the sex ratio is homogeneous among samples.

On the above evidence we conclude that almost all cases of muscular dystrophy from an affected parent (class 1) are due to one or more "dominant" genes with recognizable effects in heterozygotes of sufficient age and of either sex.

Segregation analysis of class 2 and 4(2) (limb-girdle)

Class 2 consists of cases from normal parents with at least one affected girl in the family, and class 4(2) consists of cases from normal parents with no affected girl in the family, but of the same phenotype as class 2. The probability of only one affected in a sibship of class 2, if there is at least one affected, is

$$P_{(r=1|\,r>0)} = \frac{sp\pi[x + (1 - x)q^{s-1}]}{xsp\pi + 2(1 - x)[1 - (1 - p\pi)^s - (\frac{1}{2}p + q)^s + (\frac{1}{2}p + q - \frac{1}{2}p\pi)^s]}$$

(problem 4)

and the distribution of affected individuals in sibships with more than one affected is

$$P_{(r|\,r>1)} = \frac{\binom{s}{r}p^r q^{s-r}[1 - (1 - \pi)^r][1 - (\frac{1}{2})^r]}{1 - (1 - p\pi)^s - (\frac{1}{2}p + q)^s + (\frac{1}{2}p + q - \frac{1}{2}p\pi)^s - \frac{1}{2}\pi spq^{s-1}}$$

(problem 5)

where r is the number of affected sibs including at least one affected girl and s is the sibship size (Morton, 1958, 1959). In class 4(2) with females excluded, the corresponding probabilities are the same as for problems 2 and 3.

The proportion x of sporadic cases may be estimated from the inbreeding coefficients of n isolated probands (F_I) and N familial probands (F_F) as

$$x^* = cy = c(F_F - F_I)/F_F - \alpha)$$

where α is the mean inbreeding coefficient in the general population and c = n/(n + N), and y is the proportion of sporadic cases among isolated cases, estimated from the relation $F_I = y\alpha + (1 - y)F_F$. The large-sample estimate of the variance of x^* is $\sigma^2_{x^*} = c^2\sigma^2_y + y^2\sigma^2_c$, where $\sigma^2_c = nN/(n + N)^3$ and

$$\sigma^2_y = \frac{\sigma^2_{F_F}(F_I - \alpha)^2}{(F_F - \alpha)^4} + \frac{\sigma^2_{F_I}}{(F_F - \alpha)^2}$$

(Chung, Robison, and Morton, 1959). These estimates may be converted to scores for x by the transform $K_{xx} = 1/\sigma^2_{x^*}$, $U_x = (x^* - x)K_{xx}$. Weighting each sample by the number of probands, we find $\alpha = 62 \times 10^{-5}$, $F_I = 119 \times 10^{-5}$, $F_F = 804 \times 10^{-5}$, and $x^* = .500 \pm .076$.

TABLE 2. SEGREGATION ANALYSIS OF CLASS 2 AND 4(2)

Source	Problem No.	p	x	r	U_p	U_x	U_r	K_{pp}	K_{xx}	K_{rr}	K_{px}	K_{pr}	K_{xr}
Class 2													
a > 0	4	.206	.400	.528	21.404	−18.915	−5.814	544.47	233.07	22.85	−315.00	−100.00	72.93
a > 0	5	.206	—	.528	38.593	—	−1.237	1161.40	—	4.16	—	−67.42	—
a > 0	9	—	—	.528	—	—	−58.722	—	—	260.17	—	—	—
a > 0	6	—	—	.528	—	—	9.051	—	—	86.96	—	—	—
Inbreeding	—	—	—	—	—	17.375	—	—	173.75	—	—	—	—
Total					59.997	−1.540	−56.722	1705.87	406.82	374.14	−315.00	−167.42	72.93
Sex of affected	2	½	0	1	−29.240	—	—	260.37	—	—	—	—	—
Sex of affected	3	½	—	1	−13.910	—	—	95.62	—	—	—	—	—
Sex of normal	10	½	—	—	−4.000	—	—	2184.00	—	—	—	—	—
Class 4(2) Males Only													
♂, a > 0	2	.206	.400	.528	−3.892	3.084	0.951	378.26	77.11	8.17	−165.97	−54.68	25.05
♂, a > 0	3	.206	—	.528	2.211	—	−0.040	92.80	—	0.49	—	−6.71	—
♂, a > 0	9	—	—	.528	—	—	−10.655	—	—	61.70	—	—	—
♂, a > 0	6	—	—	.528	—	—	35.213	—	—	129.24	—	—	—
Inbreeding	—	—	—	—	—	−0.633	—	—	2.04	—	—	—	—
Total					−1.681	2.451	25.469	471.06	79.15	199.60	−165.97	−61.39	25.05

	d.f.	x^2	P
p between 2 and 4(2)	1	0.55	.5–.3
x between 2 and 4(2)	1	0.08	.8–.7
Sex of affected, p = ½	1	5.23	.05–.02
2 + 4(2) Pooled			
Deviation from p = .206	1	1.59	.3–.2
p among sources	3	0.65	.9–.8
x between inbreeding and segregation data	1	2.40	.2–.1
p among samples	6	9.25	.2–.1
x among samples	6	11.14	.10–.05

From the ages in class 2 sibships, the average penetrance is calculated to be .794 for affected individuals who are not probands, .810 for all affected individuals, and .836 for their normal sibs, or approximately .825 for all sibs, giving p = .206 on the hypothesis of recessive inheritance for familial and chance isolated cases if there is complete penetrance in individuals of sufficient age.

The segregation analysis based on preliminary estimates of p = .206, x = .400, π = .528 is shown in table 2. The information matrix for x and π inverts to the covariance matrix

$$\begin{bmatrix} 21.3112 & -3.6394 \\ -3.6394 & 18.0510 \end{bmatrix} 10^{-4},$$

yielding

$$x = .400 + (0.911)(21.3112)10^{-4} + (-31.253)(-3.6394)10^{-4} = .413$$

$$\pi = .528 + (0.911)(-3.6394)10^{-4} + (-31.253)(18.0510)10^{-4} = .471$$

At these values the estimate of U_p by Taylor's theorem is

$$U_p = 58.316 + (.400 - .413)K_{px} + (.528 - .471)K_{p\pi} = 51.527.$$

The information matrix for p, x, and π inverts to the covariance matrix, the first element of which is $\sigma_p^2 = 5.9897 \times 10^{-4}$, giving $\chi_1^2 = (51.527)^2(5.9897)10^{-4} = 1.59$ as the test of the hypothesis that p = .206.

The data indicate that about 41 per cent of group 2 and 4(2) cases are sporadic, this fraction being homogeneous between inbreeding and segregation estimates and between classes. Estimates from different samples are consistent, indicating that all investigators have included similar proportions of sporadic cases. In segregating families, which give rise to the familial and chance isolated cases, the inheritance is monomeric recessive with complete penetrance in individuals of sufficient age. The segregation analysis gives no evidence on whether these genes in different families are allelic or at different loci. Failure of previous investigators to demonstrate the sporadic nature of many of these cases is due to the crude methods of analysis available to human geneticists in the past.

The distribution of affected girls among familial cases is a truncated binomial, which may be analysed by problems 2 and 3 with x = 0, π = 1. The scores for x are $U_x = 8.667$, $K_{xx} = 401.86$, indicating that sporadic expression of sex-linked dystrophy in females or pseudo-females is not an appreciable source of bias. Nevertheless, there is a deficiency of affected girls ($\chi_1^2 = 5.23$), the estimated proportion of .38 being surprisingly homogeneous among samples ($\chi_6^2 = 0.78$, P > .99). Since there is no evidence of a corresponding excess of normal girls or of sex-biased manifestation (Chung and Morton, 1959), the deficiency of affected girls may well be due to chance.

Segregation analysis of class 3 and 4(3) (Duchenne)

Class 3 consists of affected boys from normal parents with a maternal history of affected males, consistent with sex linkage, and class 4(3) consists of affected boys from normal parents with a negative maternal history, but of the same phenotype as

class 3. The cases of Stephens and Tyler (1951), although not reported in detail, have all been classified as 3 or 4(3), since they are Duchenne in type and include only affected boys. Because of the transmission through carrier females to their sons, this genetic type has been considered to be due either to a recessive sex-linked gene, or to an autosomal gene that is dominant in males but recessive or not expressed in females. We shall consider the analysis under sex-linkage first, and then demonstrate that autosomal inheritance can be excluded.

Deleterious sex-linked genes are subject to stringent selection in males, one-half of the genes transmitted from carrier mothers being exposed to elimination in this way each generation. Therefore an appreciable fraction of affected cases should result from mutation in this or the preceding generation, and appear in class 4(3) as isolated or familial cases without affected relatives. If u and v are the mutation rates and P and Q the gene frequencies in eggs and X-bearing sperm, respectively, and m the coefficient of selection against the trait, the loss of genes by selection will be $m(Q + PQ)$ and the gain by mutation $2u(1 - P) + v(1 - Q)$. At equilibrium the gain and loss are equal, and if $m \gg 0$, we may neglect the product terms PQ, uP, and vQ in comparison with P, Q, u, and v, and write $Q = (2u + v)/m$. The probability that an affected male in the general population be sporadic is $x = u/Q = mu/(2u + v)$, assuming that all cases are sex-linked. Similarly, $Q = P + u$, or $P = (v + 2u - mu)/m$. The probability that a carrier female acquire the gene through an affected father is $Q(1 - m)$, and by mutation is $u + v$, and in both these cases she will have no affected brothers. But $Q(1 - m) + u + v = P$, and since the frequency of carrier females is 2P, the probability that the brother of a carrier female not be at risk is $x' = P/2P = 1/2$.

If we consider that the mother of two or more affected boys is a carrier, the probability that the grandmother was a carrier is $1 - x'$. Assuming ascertainment through nephews, the probability that none of s maternal uncles be affected is $x' + (1 + x')q^s$, counting a set of maternal uncles once for each index sibship with two or more boys affected (problem 7).

If the index case has no affected siblings, and the number of boys in the sibship is n, the probability that he is sporadic is $x/\{x + (1 - x)q^{n-1}\}$, and the probability that none of s maternal uncles be affected is

$$P(r = 0) = \{x + (1 - x)q^{n-1}[x' + (1 - x')q^s]\}/\{x + (1 - x)q^{n-1}\},$$

the complementary probability of at least one affected uncle being

$$P_{(r>0)} = \frac{(1 - x)(1 - x')q^{n-1}(1 - q^s)}{x + (1 - x)q^{n-1}} \qquad \text{(problem 8)}$$

where r is the number of affected maternal uncles. If no maternal uncle is affected, the maximum likelihood scores for x and x′ are

$$u_x = \{B - C + q^{n-1}(C - AB)\}/BC \qquad (x, x' > 0)$$

$$u_{x'} = (1 - x)q^{n-1}(1 - q^s)/C$$

where

$$A = x' + (1 - x')q^s$$

$$B = x + (1 - x)q^{n-1}$$

$$C = x + (1 - x)q^{n-1}A.$$

If at least one uncle is affected,

$$u_x = -1/B(1 - x) \qquad (x, x' > 0)$$

$$u_{x'} = -1/(1 - x').$$

The variances and covariance are $k_{xx} = E(u_x^2)$, $k_{x'x'} = E(u_{x'}^2)$ and $k_{xx'} = E(u_x U_{x'})$.

Isolated and familial cases without regard to maternal uncles, maternal uncles of familial cases, and maternal uncles of isolated cases constitute three independent sources of information about x and x′. Agreement among them, and of x′ with 1/2, lends support to any conclusion about the deviation of x from m/3, and hence the equality of the mutation rates in the two sexes.

Since individuals with this type of muscular dystrophy rarely reproduce, m is nearly 1 and x approaches ⅓ on the hypothesis of equal mutation rates in the two sexes. From the ages of males in class 3 sibships, the average penetrance is .869 for affected individuals who are not probands, .866 for all affected individuals, and .839 for all sibs, giving p = .427 on the hypothesis of sex-linkage.

The segregation analysis, based on these values and a preliminary estimate of $\pi = .519$, is shown in table 3. An improved estimate of π is $.519 + U_\pi/K_{\pi\pi} = .492$. The information matrix for p and π inverts to the covariance matrix, the first element of which is $\sigma_p^2 = 8.8640 \times 10^{-4}$. By Taylor's theorem the score for p at $\pi = .492$ is $U_p = -3.488 + (.519 - .492)(106.60) = -0.610$, giving

$$\chi_1^2 = (-.610)^2(8.8640)10^{-4} = .0003.$$

Similarly, from the covariance matrix for x and π, the test for the hypothesis that $x = ⅓$ is $\chi_1^2 = .216$. Since there is no covariance between x′ and π, the corresponding test for the hypothesis that $x' = ½$ is simply $\chi_1^2 = (U_{x'})^2/K_{x'x'} = 1.72$. The estimates of p, x, and x′ are homogeneous among sources and among samples with the exception of x among samples. Walton's sample (1955) includes an unusually small frequency of sporadic cases, and the heterogeneity generated by this discrepancy is barely significant ($\chi_7^2 = 14.08$, P = .048). Considering the small amount of information in the individual sources and the approximate nature of the χ^2 test in small samples, the biological significance of this heterogeneity is doubtful. We conclude that the data are in agreement with the hypothesis that dystrophy in class 3 and 4(3) is due to a sex-linked recessive gene or genes with complete penetrance in males of sufficient age and with equal mutation rates in eggs and sperm. It is impossible to prove the assumption that all isolated cases are sex-linked since affected males rarely reproduce, but the agreement with sex-linkage is remarkable and we have shown that

TABLE 3. SEGREGATION ANALYSIS OF CLASS 3 AND 4(3)—MALES ONLY

Source	Problem No.	P	x	x′	r	U_p	U_x	$U_{x'}$	U_r	K_{pp}	K_{xx}	$K_{x'x'}$	K_{rr}	K_{px}	K_{pr}	K_{xr}	$K_{xx'}$
a > 0	2	.427	⅓	—	.519	5.590	−6.670	—	−1.826	311.78	306.27	—	25.66	−268.98	−80.10	88.50	—
a = 0	2	.427	0	—	1	−28.110	—	—	—	413.55	—	—	—	—	—	—	—
a > 0	3	.427	—	—	.519	.440	—	—	−.178	271.87	—	—	2.69	—	−26.50	—	—
a = 0	3	.427	—	—	1	18.592	—	—	—	150.96	—	—	—	—	—	—	—
a > 0	9	—	—	—	.519	—	—	—	−21.870	—	—	—	201.38	—	—	—	—
a > 0	6	—	—	—	.519	—	—	—	8.377	—	—	—	338.32	—	—	—	—
Uncles	7*	.427	—	½	—	—	—	4.249	—	—	—	92.72	—	—	—	—	—
Uncles	8*	.427	⅓	½	—	—	13.646	14.550	—	—	113.42	112.97	—	—	—	—	109.80
Total						−3.488	6.976	18.799	−15.497	1148.16	419.69	205.69	568.05	−268.98	−106.60	88.50	109.80

	d.f.	x^2	P
Deviation from p = .427	1	.0003	.99–.98
Deviation from x = ⅓	1	0.22	.7–.5
Deviation from x′ = ½	1	1.72	.2–.1
p among sources	3	4.29	.3–.2
p among samples	8	11.42	.2–.1
x between sources	1	1.67	.2–.1
x among samples	7	14.08	.05
x′ between sources	1	0.35	.7–.5
x′ among samples	6	9.08	.2–.1

* Excluding Levison's sample, which does not report maternal uncles of class 4(3).

the frequency of misdiagnosis is small and tends to be compensated by misclassification of sex-linked cases as limb-girdle dystrophy (Chung and Morton, 1959).

Haldane (1956) first examined the question of equality of mutation rates in the two sexes. He made two assumptions that tended to overestimate x, namely that $p = \frac{1}{2}$ after an attempted age correction and that x is the same as the proportion w of families with sporadic cases among families with affected members, whereas in fact $w = xsp/\{1 - q^s - x(1 - sp - q^s)\}$ (Morton, 1959). He also made two assumptions which tended to underestimate x, namely that $\pi = 1$ and that sibships without at least one proband could be treated in the same way as index sibships. Cheeseman et al. (1958) and Smith and Kilpatrick (1958) corrected two of these sources of bias, reversing Haldane's conclusion that there was a significant deficiency of sporadic cases. They retained his assumption that $p = \frac{1}{2}$, $\pi = 1$, however. Our analysis removes these restrictions and incorporates the evidence from several other studies and from uncles of isolated and familial cases, applying the methodology of Morton (1958). Our estimate of $x = .355 \pm .050$, although with a smaller standard error, is very similar to the estimate of Smith and Kilpatrick of $.351 \pm .111$.

Several lines of evidence argue against the theory of an autosomal gene that is dominant in males and recessive or not expressed in females. On that hypothesis, half of the sons of affected males should also be affected, while none should be affected under sex-linkage, barring the highly unlikely events of mutation or mating to a carrier. In the pooled material there were 10 sons of class 3 cases and 8 sons of class 4(3) cases, all normal. (The 9 daughters of affected individuals were also normal, although several transmitted the gene to sons or grandsons, as would be expected under either hypothesis.) Walton (1957) reported a case of Turner's syndrome (ovarian agenesis) with Duchenne muscular dystrophy in a sex-linked pedigree. If the gene were autosomal, this would require a coincidence of two rare events (sexual maldevelopment and heterozygote expression or mutation). However, if the gene were sex-linked, only one unusual event is required, since many of the cases with Turner's syndrome born to a carrier mother are expected to have muscular dystrophy, being XO in chromosome type (Ford et al., 1959). Philip and Walton (1956) presented evidence consistent with linkage between this type of dystrophy and color-blindness, although far from proving such linkage. An attempt to find evidence for autosomal linkage was unsuccessful (Walton, 1955).

Linkage is a less powerful test of the hypothesis of sex-influenced dominance than the following method, which is applicable even when affected individuals fail entirely to reproduce. For such a gene, being autosomal, the gene frequencies in males and females must be the same, say P. Then the gene frequency in sperm is $P(1 - m) + u(1 - P)$, which approaches $u + P(1 - m)$, and the gene frequency in eggs is $P + v(1 - P)$, which approaches $v + P$. The genotype frequency in both sexes is $2P = u + P(1 - m) + v + P$, or $P = (u + v)/m$. The probability that the grandfather was affected and transmitted the gene to the mother is $P(1 - m)$. The probability that the grandmother was heterozygous and transmitted the gene to the mother is P. The probability that the mother is a carrier by mutation is $u + v$. Therefore the probability that the brother of a carrier female not be at risk is $x' = (u + v)/(2P - Pm + u + v) = m/2$, which is not in practice distinguishable

from the relation $x' = \frac{1}{2}$ for sex-linked genes. The probability of an affected male by mutation is $u + v$, and by inheritance from the mother is P. If the father is normal, the frequency of mutations among affected cases is

$$x = (u + v)/(P + u + v) = m/(m + 1),$$

which approaches $\frac{1}{2}$ for a highly deleterious gene, whereas x approaches $\frac{1}{3}$ for a highly deleterious sex-linked gene if the mutation rates in eggs and sperm are equal. By Taylor's theorem, the score for x at $\pi = .492$ is 9.366. From the covariance matrix for x and π, $\sigma_x^2 = 24.6365 \times 10^{-4}$. The estimate of x from the data is $.333 + U_x/K_{xx} = .355$, giving $\chi_1^2 = (.500 - .355)^2/\sigma_x^2 = 8.53$, which is a highly significant deviation from the expectation under sex-influenced dominance in the direction expected under sex-linkage. We conclude that the locus for this type of dystrophy is on the X-chromosome.

Ascertainment and incidence

Although the estimates of p, x and x′ are consistent among sources and problems, the estimates of π are highly heterogeneous. None of the previous studies defined the terms *propositus*, *proband*, or *ascertainment*, and it is apparent that their usages were not always consistent or correct. Stephens and Tyler (1951) reported an incidence in Utah that has never been exceeded in any other sample, indicating relatively complete ascertainment, yet except for a pair of affected twins they distinguished only one proband in each family. Walton (1955), reporting a smaller incidence of dystrophy, enumerates such a large number of ascertainments that, if they were independent, selection of cases must have been virtually exhaustive. Stevenson, after a careful study of the cases in Northern Ireland (1953), has continued to ascertain additional cases in the same population (1955, 1958). From our own experience in Wisconsin, we doubt that exhaustive ascertainment is possible for a condition not requiring institutionalization and not always correctly diagnosed.

In class 1, the estimate of π from ascertainments in Walton's data is .497, while in all other samples it is only .075. The Wisconsin material gives $\pi = .235$, which is very close to the mean for all samples. It seems likely that over-reporting of ascertainments in Walton's study is compensated by under-reporting of probands in other instances. Since the covariance between p and π is almost negligible in large pedigrees, the segregation analysis of class 1 is not greatly affected by inaccuracy in the pooled estimate of π.

The remaining classes give homogeneous estimates of π when compared within samples. However, the samples are significantly heterogeneous, even when Walton's study is excluded. For his data $\pi = .90$, for the Wisconsin material $\pi = .39$, and for the other samples $\pi = .31$. It seems likely that the mean value of π is greater than for the Wisconsin sample, since we have included only sibships from which at least one affected member underwent examination at the Muscular Dystrophy Clinic, thus excluding sibships that were represented only by patients who were unwilling or unable to make this trip. The pooled estimates of $\pi = .471$ for class 2 and 4(2) and $\pi = .492$ for class 3 and 4(3) therefore seem reasonable.

TABLE 4. SUMMARY OF THE SEGREGATION ANALYSIS

Class	Parameter	Expected	Observed
1	p	.402	.420 ± .019
2 + 4(2)	p	.206	.230 ± .024
	x	—	.413 ± .046
3 + 4(3)	p	.427	.42 ± .030
	x	.333	.355 ± .050
	x′	.500	.591 ± .070

The Wisconsin data are the most critical for estimation of π. A *proband* was defined as an affected person referred by at least one physician to the Muscular Dystrophy Clinic and examined there, independently of other cases. Many of these patients were referred without previous inquiry by the clinic, others who had been previous inpatients in the University Hospitals were re-contacted through their referring physician, and still others were ascertained through the cooperation of the State Bureau for Handicapped Children in contacting the local physician. Finally, a descriptive letter and questionnaire was sent to all members of the State Medical Society, advising them of the functions of the clinic and indicating interest in examining cases of dystrophy. No attempt was made to interview patients who were unwilling to be examined. Nearly all patients referred in these ways were extremely cooperative and have been seen several times in the clinic. Presumably the less interested patients did not consent to referral. All affected family members mentioned in a referral who subsequently came to the clinic were considered probands, whether or not they accompanied the index case on his first appointment. Siblings who became affected after examination of the index case were not considered probands.

Each referral by a different physician constituted an ascertainment. Thus if an individual was referred by Drs. X, Y, and Z, this was treated as three ascertainments regardless of the fact that Dr. X may have notified the clinic of this patient several times, through the University Hospitals, the Bureau for Handicapped Children, and in response to the letter sent to members of the State Medical Society. The high frequency of ascertainments per proband reported by Walton (1955) is probably due multiple counts of such dependent notifications, which do not follow the truncated Poisson distribution of problem 6.

Using these definitions, the scores for classes 2, 3, 4(2), and 4(3) are $U_\pi = -56.370$, $K_{\pi\pi} = 283.50$ for problem 6, and $U_\pi = 10.140$, $K_{\pi\pi} = 66.44$ for problem 9. Much the greater amount of information is supplied by the number of ascertainments (problem 6), but the two estimates are heterogeneous ($\chi^2_1 = 6.65$). We suspect that problem 6 may give too small an estimate ($\pi = .32$) because referral by one physician tends to make referral by another less likely, while problem 9 may give too large an estimate ($\pi = .67$) because two or more affected relatives may be examined and counted as probands, even when only one of them would have submitted to examination independently. For example, early cases are more likely to be brought to the clinic than terminal ones, but a terminal case may accompany an affected younger sibling to the clinic and be counted as a proband. Considering all the possible biases

to which any estimate of ascertainment is exposed, the agreement between the two methods is reasonably good. We have more confidence in the pooled estimate of $\pi = .39$ than in either estimate separately.

Although π can hardly have been the same for all these studies, discrepancies among them and errors in determining the pooled value do not seem to have affected the segregation analysis, which gives internally consistent results remarkably close to expectation. In part this must be because of compensation of errors in estimating π, but also the covariances between π and the other parameters are small enough so that the analysis is rather robust, not appreciably disturbed by small departures from the ascertainment model. Since even the incomplete and crude data on ascertainment provided by these muscular dystrophy studies can apparently be tolerated, an enlightened effort to define probands and ascertainments appropriately seems all that is required to assure a valid segregation analysis by these methods.

The most serious effect of errors in estimating π is on the calculation of incidence. This has not been appreciated in human genetics because the ascertainment probability is almost never taken into account in estimating incidence, hence the recommendation by some authors of single selection, which gives no estimte of π, because it simplifies the arithmetic of segregation analysis.

Let $P = n/\pi N$ be the prevalence of a trait in the general population, where n is the number of living probands and N is the population size. To obtain I, the incidence at birth (and presumably at conception) of individuals who will develop the trait, N should be replaced by $N' = \int_z n(z)P(z)\,dz$, where n(z) is the number in the general population at age z and P(z) is the probability that a person capable of developing the trait do so by age z and not die of it before then. For practical computation P(z) may be replaced approximately by $G(z)[1 - D(z)]$, where G(z) is the cumulative frequency of onset at age z and D(z) is the cumulative frequency of death at age z among affected individuals. These distributions are given by Chung and Morton (1959); n(z) has been taken from the Demographic Yearbook for 1955.

In class 1 there are 30 living probands in the pooled sources, drawn from a total population of about 23,488,000. The pedigree of Tyler and Stephens (1950) has been omitted because the number of probands and other data were not reported. The pooled estimate of π is .229, from which the prevalence is calculated as $P = 5.6 \times 10^{-6}$. We find $N' = .606N$, from which the incidence at birth of persons who will develop the disease is $I = 9.2 \times 10^{-6}$. In the Wisconsin sample, $n = 2$, $\pi = .235$, $N = 3{,}700{,}000$, and the corresponding estimates are $P = 2.3 \times 10^{-6}$, $I = 3.8 \times 10^{-6}$.

These calculations are subject to two principal errors. Biases in n and π tend to be compensatory, but the frequency of a dominant trait that is far from lethal may be expected to vary considerably among populations, sometimes producing enormous pedigrees like the one reported by Tyler and Stephens.

In class 2 and 4(2) there are 267 living probands, $\pi = .471$, and $N' = .304N$. We find $P = 24.1 \times 10^{-6}$, $I = 79.4 \times 10^{-6}$. For Wisconsin, $n = 26$, $\pi = .354$, $P = 19.9 \times 10^{-6}$, $I = 65.3 \times 10^{-6}$.

In class 3 and 4(3), $n = 247$, $\pi = .492$, $N' = .236N$, and the number of males in the population is 11,998,000, including the sample of Stephens and Tyler (1951).

The estimate of prevalence is $P = 41.8 \times 10^{-6}$, and $I = 177.3 \times 10^{-6}$ among males. For Wisconsin males, $n = 51$, $\pi = .419$, $P = 65.8 \times 10^{-6}$, $I = 278.8 \times 10^{-6}$.

Pooling classes 1, 2, 3, and 4, the total frequency of dystrophy per million persons is $P = 50.6$, $I = 177.2$ for the pooled sample, and $P = 55.1$, $I = 208.5$ for Wisconsin.

In class 1 the gene frequency at birth is $q = I/2$, or 2×10^{-6} based on the Wisconsin sample. In class 3 and 4(3), the gene frequency in males is $q = I$, or 279×10^{-6}. By enumerating cases born during an extended period, Walton (1955), Stevenson (1958) and Stephens and Tyler (1951) obtained minimal estimates of 130, 182, and 286×10^{-6}, respectively.

If more than one locus contributes to class 1 or class 3 and 4(3), q is the sum of the gene frequencies at these loci, since I is linear in q. The problem is more delicate for class 2 and 4(2), since with recessive genes I is quadratic in q. By the theory of detrimental equivalents (Chung, Robison, and Morton, 1959), the incidence is $I = A + B\alpha$ and the mean inbreeding coefficient of probands is $\bar{F} = [A\alpha + (\sigma^2 + \alpha^2)B]/I$, where A is the incidence in a random mating population, α, σ^2 are the mean and variance of the inbreeding coefficient in the general population, and B is the mean number of detrimental equivalents per gamete. For genes with complete penetrance in homozygotes of sufficient age, the total gene frequency per gamete lies between B and A + B. These quantities are estimated as

$$A = I - B\alpha$$

$$B = I(\bar{F} - \alpha)/\sigma^2$$

For recessive cases we find $\bar{F} = 804 \times 10^{-5}$, $I = (.587)(65.3 \times 10^{-6}) = 38 \times 10^{-6}$, $\alpha = 62 \times 10^{-5}$, and $\sigma^2 = 35 \times 10^{-6}$, weighting each sample by the number of probands.

$$A = (33 \pm 2) \times 10^{-6}$$

$$B = .0081 \pm .0027$$

A minimum estimate of the number of loci contributing to recessive cases of dystrophy, obtained on the assumption of equal gene frequencies, is $n = B^2/A = 2.0 \pm 1.3$. Most familial and chance isolated cases of class 2 and 4(2) are evidently due to genes at a few loci, perhaps at a single locus. The average gene frequency per locus is $q = A/B = .0041 \pm .0014$. If all sporadic cases were heterozygotes for one of these genes, the penetrance in heterozygotes would be $h = (.413)(65.3 \times 10^{-6})/2B = .0017$. That is, if 17 out of 10,000 carriers developed dystrophy because of exogenous agents or genetic factors at other loci, this would account for the observed frequency of sporadic cases.

If the load represented by B were segregational (Crow, 1958), and the gene frequencies maintained by heterozygote advantage over dystrophic homozygotes, there would have to be more than $B/A = 244$ alleles at each locus capable of producing dystrophy, each allele giving a superior heterozygote with all the others! This is more than 100 times the number of genes indicated by the data. The frequency of each allele and the selection coefficient against the normal homozygote must be less than $B/244 = 3.3 \times 10^{-5}$, which is so small that virtually all homo-

zygotes would come from consanguineous marriages. Since the postulated genes are individually very rare, the advantage of heterozygotes inter se would have to be several orders of magnitude less than this, and therefore negligible by comparison with the mutation pressure. It is incredible that so many rare alleles could be maintained by selection in a finite population. We conclude that the load represented by B is mutational, as would be expected for rare and presumably hypomorphic genes.

The argument in the above paragraph depends on the assumption that homozygotes are eliminated as dystrophic cases, and does not apply if selection is partly against some other homozygous effect. However, on the hypothesis of heterozygote advantage, the selection exerted at each locus capable of producing dystrophy is at least qm, where m is the selection coefficient against dystrophic homozygotes. Since $q = .0041$ and $m = .746$ (see below), each dystrophy locus is responsible for .003 eliminations per zygote. If we make the liberal supposition that the total number of expressed lethal equivalents is .60, so that each generation is produced by $e^{-0.6} = 55$ per cent of the previous generation, then 200 loci with mutants as detrimental as the ones that produce muscular dystrophy are sufficient to account for all selection, both genetic and phenotypic. Since the number of loci with major effects is of the order of 10,000 per gamete in Drosophila, and presumably at least as many in man, it must be true that the great majority of detrimental alleles are maintained by some other mechanism than heterozygote advantage, and the load represented by detrimental equivalents is largely mutational.

Fitness and mutation

At least three approaches have been made to the estimation of fitness for specific traits in man. The simplest is to take the ratio between the mean number of children born to affected individuals and to a suitable control sample, both of which have terminated their reproduction. Krooth (1955) has proposed a modification of this based on sib controls. Unfortunately, sib controls are neither feasible nor desirable: not feasible, because many sibs will not have completed their reproduction and their total progeny will not be known; not desirable, because information is lost about possible differences between the reproductive performance of sibs and the general population. Finally, Reed and Neel (1955) have proposed a measure of relative reproductive span, based on the assumption that fecundity up to the time of death is the same as for the general population. This method is clearly not applicable to diseases like muscular dystrophy.

To adjust for possible ascertainment bias due to association between number of sibs and number of children (Krooth's problem), let $s - 1$ be the number of sibs of an affected person and n be the number of his children. The regression of n on $s - 1$ is $n = a + b(s - 1)$. In the general population the mean value of $s - 1$ is $\mu = \Sigma s(s - 1)/\Sigma s = \bar{s} - 1 + \sigma^2/\bar{s}$, where $\bar{s}$ and σ^2 are the mean and variance of s among families with $s > 0$. Typical values of μ are 3.17 for a rural Italian population (Cavalli, unpublished) and 3.64 for an American population a generation ago (Baber and Ross, 1924). If b is not significantly different from zero, the selection coefficient against affected persons may be estimated as $m = 1 - \bar{n}/\bar{N}$, where $\bar{n}$, $\bar{N}$ are the mean number of children from affected persons and a suitable control. If b is sig-

TABLE 5. FERTILITY OF DEAD CASES

Class	Number of Cases	Number of Children n	Adjusted Number of Children $a + b\mu$	Selection Coefficient m
1	40	3.975	3.142	?
2 + 4(2)	62	0.694	0.642	.746
3 + 4(3)	112	0.116	.095	.962

nificantly different from zero, the selection coefficient is $m = 1 - (a + b\mu)/\bar{N}$, adjusted for possible ascertainment bias due to familial association in fertility. In the dystrophy material the deviation of b from zero is significant in class 1, and we have therefore used the adjusted estimate.

The fecundity of the general population represented by these samples is approximately twice the ratio of the population in 1950 to the population in 1920, or $\bar{N} = 2.524$. The values of $\bar{n}$ and m are shown in table 5, using dead cases as the closest practical approach to a representative sample of reproductive performances and taking $\mu = 3.5$. The selection coefficient is calculated to be about .746 for class 2 and 4(2) and at least .962 for class 3 and 4(3). Since several of the fertile Duchenne cases were ascertained through affected grandchildren, the selection coefficient is probably greater than this. The fertility of class 1 is actually greater than the estimated population mean, probably because many cases are drawn from previous generations. Since the mean number of sibs is 5.400, the selection coefficient against these cases is evidently less than $1-3.975/5.4 = .264$.

Assuming equilibrium between selection and mutation, the mutation rate for dominant genes is not more than $u = mq = 5.0 \times 10^{-7}$ for class 1. Therefore the contribution of dominant mutations to sporadic cases of class 2 and 4(2) is negligible. The rate for recessive genes is $u = q(\alpha' + q + h)m$, where α' is the long-term inbreeding coefficient and h is the selection coefficient against heterozygotes. The value of α' is probably about .006 (Neel *et al.*, 1949). For $h = 0$, we find $u = 3.1 \times 10^{-5}$ per locus for the recessive genes contributing to class 2 and 4(2). If all sporadic cases are due to penetrance in heterozygotes, $h = .0017$, $u = 3.6 \times 10^{-5}$. The assumption of equal gene frequencies, which makes the estimate of the number of loci a minimum, makes this estimate of u a maximum if there is more than one locus. The mutation rate for sex-linked genes, assuming equality in eggs and sperm, is $u = mq/3 = 8.9 \times 10^{-5}$ for class 3 and 4(3). Without assuming equilibrium or equal rates in the two sexes, but using the estimated frequency of sporadic cases, the mutation rate in eggs is $u = .355q = 9.9 \times 10^{-5}$. The difference between the two estimates is trivial.

A variety of evidence indicates that, in mammals, formation of oocytes is virtually complete soon after birth (Parkes, 1956). On this hypothesis, the maternal age at birth of sex-linked mutants is critical for distinguishing two types of mutation. On one model, mutation is a failure to copy the gene correctly at cell division, in which case the mutation rate should not be related to maternal age. On the other model, mutation is caused by radiation or chemical mutagens independently of cell division and therefore accumulates with maternal age. The mean maternal age at birth of 72 isolated cases of class 4(3) is 27.736, and the mean for 147 normal sibs is 27.864, the difference being $-.128 \pm .922$ in the direction contrary to the accumulation model.

The data concur with more indirect evidence from maternal ages for autosomal dominant mutants (Penrose, 1955), in suggesting that the mutation rate in human gametes is more closely related to the number of previous cell divisions than to gonial age.

Genetic risks

In class 1, .87 of cases reported onset by age 20. Presumably nearly all carriers could be detected by a careful neurological examination at maturity. The probability that the child of a carrier be affected is ½, and the probability that the child of a sibling who is not a carrier be affected is zero. The risk for a carrier's child who is normal at age z to develop the disease is not more than $[1 - G(z)]/[2 - G(z)]$, where $G(z)$ is given in figure 3 of Chung and Morton (1959). For example, since $G(z) = .87$ at age 20, the risk for a normal 20-year-old child of a carrier to develop the disease is not more than .115, and is probably much less than this, since $G(z)$ is based on recollection rather than clinical examination.

In case 2 and 4(2), sibships with familial cases, or an isolated case with parental consanguinity, have a risk of ¼ for dystrophy with each subsequent child. The probability that an isolated case with S normal siblings be sporadic is $Q = .413 \Big/ \left\{.413 + .587 \prod_{i=1}^{S} [1 - G(z_i)/4]\right\}$, where z_i is the age of the i^{th} sib. Since the mean penetrance is $y = .836$ for unaffected sibs, this function may be approximated by $Q = .413/\{.413 + .587(.791)^S\}$, as shown in table 6.

The risk for dystrophy in the child of an affected homozygote is expected to be $P = F + q + h$, where F is the inbreeding coefficient, q the gene frequency, and h the penetrance in heterozygotes. The expected risk is P/3 for the children of normal sibs of affected homozygotes, P/2 for the children of carriers, and P/4 for the children of the sibs of carriers. The empirical risks for relatives of index cases are shown in table 7, including three families from marriages of affected homozygotes to presumptive carriers (Class 5 of Chung and Morton, 1959). The observations are consistent both with no risk for the relatives of sporadic cases and with a risk half as great as for the relatives of familial cases, as would be expected if sporadic cases are penetrant heterozygotes. The risk in relatives of familial cases is consistent with complete re-

TABLE 6. THE PROBABILITY THAT AN ISOLATED CASE OF CLASS 2 OR 4(2) BE SPORADIC, IF THERE ARE S NORMAL SIBS

S	0	2	4	6	8	12	∞
Q	.41	.53	.64	.74	.82	.92	1

TABLE 7. GENETIC RISKS FOR CLASS 2 AND 4(2) RELATIVES

	r = 1		r > 1		
	Affected	Normal	Affected	Normal	kP
Children	0	110	0	66	P
Sibs' children	0	348	1	136	P/3
Carriers' children	2	1755	3	634	P/2
Carriers' sibs' children	0	1409	0	414	P/4
P	.0014 ± .0010		.0075 ± .0037		

TABLE 8. THE UPPER LIMIT TO THE PROBABILITY THAT A BROTHER OF AN ISOLATED CASE OF CLASS 4(3) BE AFFECTED OR THAT A SISTER BE A CARRIER, IF THERE ARE S NORMAL BROTHERS AND s NORMAL MATERNAL UNCLES

	S Brothers				
s Uncles	0	2	4	6	∞
0	.33	.20	.09	.04	0
2	.28	.15	.06	.02	0
∞	.25	.13	.05	.02	0

cessivity ($P = \alpha + q = .0047$) or with occasional penetrance in heterozygotes ($P = \alpha + q + h = .0064$). All evidence indicates that the risk in noninbred relatives of class 2 and 4(2) is very small.

In class 3 and 4(3) the risk is ½ for sons of carrier women. One-half of the daughters of carrier women and all the daughters of carrier men are carriers. The probability that an isolated case with S normal brothers and s normal maternal uncles be sporadic is at least

$$Q = x \Big/ \left\{ x + (1 - x) \prod_{i=1}^{s} [1 - G(z_i)/2] \left[x' + (1 - x') \prod_{i=1}^{S} [1 - G(z_i)/2] \right] \right\}.$$

Since $x = \frac{1}{3}$, $x' = \frac{1}{2}$, and the average penetrance in normal sibs is .839, we may write approximately $Q = 1/\{1 + (.58)^s[1 + (\frac{1}{2})^s]\}$. The probability that a brother of an isolated case be affected or a sister be a carrier is not more than $(1 - Q)/2$, as given in table 8. The method for computing exact risks has been given by Binet et al. (1958), using information provided by normal male relatives more remote than uncles.

Considering the probabilities in tables 6 and 8, it is apparent that in the absence of positive identification of carriers, the separation of sporadic from chance isolated cases approaches certainty only in large families. When more is known about the biochemistry of the several genetic types of muscular dystrophy, it may be possible to determine with assurance that a given family member is, or is not, a carrier. Until then, the genetic basis, if any, of sporadic cases of class 2 and 4(2) must remain in doubt, and the specification of genetic risks be unsatisfactory.

SUMMARY

Genetic classes of muscular dystrophy delimited by pedigree information and discriminant functions were analysed by maximum likelihood scores, the theory of detrimental equivalents, and other methods.

One group (facioscapulohumeral) is due to an autosomal dominant gene or genes with complete penetrance in individuals of both sexes who survive to the age of onset. The prevalence is variable among populations, but is approximately 2 living cases per million, or an incidence of 4 persons who will develop the trait per million births. The fertility is nearly normal, and the mutation rate is not more than 5 per ten million gametes.

A second group (limb-girdle) is composed of autosomal recessive cases with penetrance in both sexes. The prevalence is 12 living cases per million, or an incidence of

38 persons who will develop the disease per million births. The fertility is 25% of normal. Most of these cases are due to a few loci, perhaps to a single locus, with a mutation rate of 31 per million loci. The frequency of heterozygous carriers, who are rarely if ever affected, is 16 per thousand persons, and this load is shown to be mutational in origin and not due to heterozygote advantage.

A third group (limb-girdle) is composed of truly sporadic cases, of unknown etiology. Fertility and other characteristics are the same as for recessive cases. The prevalence is 8 living cases per million, or an incidence of 27 persons who will develop the disease per million births. If these cases, whose etiology may well be complex, were entirely due to occasional penetrance in heterozygous carriers of recessive genes for dystrophy, the penetrance would be only 17 per ten thousand carriers. These sporadic cases are not due to dominant mutations, since all of 110 children of isolated index cases were normal.

The fourth group (Duchenne) consists of sex-linked cases, with a prevalence of 66 living cases per million males, or an incidence of 279 boys who will develop the disease per million male births. Their fertility is less than 4% of normal. One third of these cases are truly sporadic, due to new mutations. The mutation rate is 89 per million gametes, and is the same in eggs and sperm. The absence of a significantly elevated maternal age for isolated cases suggests that the mutation rate in eggs is more closely related to the number of previous cell divisions than to gonial age.

Genetic and empirical risks based on relationship, age, and the number of normal relatives are given for the different types of dystrophy.

ACKNOWLEDGMENTS

The authors wish to acknowledge the help given them by Dr. H. A. Peters and the staff of the Muscular Dystrophy Clinic in studying these families, by the Wisconsin State Medical Society and the Bureau for Handicapped Children in finding cases, and by the patients and their families whose cooperation made this study possible.

REFERENCES

BABER, R. E., AND E. A. ROSS. 1924. Changes in the size of American families in one generation. *Univ. Wisconsin Stud. in Soc. Sci. and Hist.* No. 10.

BECKER, P. E. 1933. *Dystrophia musculorum progressiva.* Thieme: Stuttgart.

BELL, JULIA. 1943. On pseudohypertrophic and allied types of progressive muscular dystrophy. *Treas. Human Inheritance* 4: 283–341.

BINET, F. E., R. J. SAWERS, AND G. S. WATSON. 1958. Heredity counselling for sex-linked recessive deficiency diseases. *Ann. Human Genet.* 22: 144–152.

CHEESEMAN, E. A., S. J. KILPATRICK, A. C. STEVENSON, AND C. A. B. SMITH. 1958. The sex ratio of mutation rates of sex-linked recessive genes in man with particular reference to Duchenne type muscular dystrophy. *Ann. Human Genet.* 22: 235–243.

CHUNG, C. S., AND N. E. MORTON. 1959. Discrimination of genetic entities in muscular dystrophy. *Am. J. Human Genet.* 11: 339–359.

CHUNG, C. S., O. W. ROBISON, AND N. E. MORTON. 1959. A note on deaf mutism. *Ann. Human Genet.* (in press).

CROW, J. F. 1958. Some possibilities for measuring selection intensities in man. *Human Biol.* 30: 1–13.

DE GROUCHY, J. 1953. *Contribution a l'étude de l'hérédité de la myopathie.* Thesis: Paris.

FORD, C. E., P. E. POLANI, J. H. BRIGGS, AND P. M. F. BISHOP. 1959. A presumptive human xxy/xx mosaic. *Nature* 183: 1030–1032.

HALDANE, J. B. S. 1956. Mutation in the sex-linked recessive type of muscular dystrophy. A possible sex difference. *Ann. Human Genet.* 20: 344–347.

KROOTH, R. S. 1955. The use of the fertilities of affected individuals and their unaffected sibs in the estimation of fitness. *Am. J. Human Genet.* 7: 325–360.

LAMY, M., AND J. DE GROUCHY. 1954. L'hérédité de la myopathie (formes basses). *J. Génét. Humaine* 3: 219–261.

LEVISON, H. 1951. Dystrophia musculorum progressiva. *Op. ex domo Biol. Hered. Human. Univ. Hafn.* vol. 26.

MORTON, N. E. 1958. Segregation analysis in human genetics. *Science* 127: 79–80.

MORTON, N. E. 1959. Genetic tests under incomplete ascertainment. *Am. J. Human Genet.* 11: 1–16.

NEEL, J. V., M. KODANI, R. BREWER, AND R. C. ANDERSON. 1949. The incidence of consanguineous matings in Japan. *Am. J. Human Genet.* 1: 156–178.

PARKES, A. S. 1956. *Marshall's physiology of reproduction.* v. 1, pt. 1. 688 pp. N. Y.: Longmans.

PENROSE, L. S. 1955. Parental age and mutation. *Lancet.* 2: 312–313.

PHILIP, URSULA, AND J. N. WALTON. 1956. Colour blindness and the Duchenne-type muscular dystrophy. *Ann. Human Genet.* 21: 155–158.

REED, T. E., AND J. V. NEEL. 1955. A genetic study of multiple polyposis of the colon (with an appendix deriving a method of estimating relative fitness). *Am. J. Human Genet.* 7: 236–263.

SJOVALL, B. 1936. Dystrophia musculorum progressiva. *Acta Psychiat. neur.*, suppl. 10.

SMITH, C. A. B., AND S. J. KILPATRICK. 1958. Estimates of the sex ratio of mutation rates in sex-linked conditions by the method of maximum likelihood. *Ann. Human Genet.* 22: 244–249.

STEPHENS, F. E., AND F. H. TYLER. 1951. Studies in disorders of muscle. V. The inheritance of childhood progressive muscular dystrophy in 33 kindreds. *Am. J. Human Genet.* 3: 111–125.

STEVENSON, A. C. 1953. Muscular dystrophy in Northern Ireland. I. An account of 51 families. *Ann. Eugen.*, Lond., 18: 50–93.

STEVENSON, A. C. 1955. Muscular dystrophy in Northern Ireland. II. An Account of nine additional families. *Ann. Human Genet.* 19: 159–164.

STEVENSON, A. C. 1958. Muscular dystrophy in Northern Ireland. IV. Some additional data. *Ann. Human Genet.* 22: 231–234.

TYLER, F. H., AND F. E. STEPHENS. 1950. Studies in disorders of muscle. II. Clinical manifestations and inheritance of facioscapulohumeral dystrophy in a large family. *Ann. Int. M.* 32: 640–660.

WALTON, J. N. 1955. On the inheritance of muscular dystrophy. *Ann. Human Genet.* 20: 1–38.

A Genetically Determined Human Serum Factor Detected by its Effect on a Mating Reaction in Yeast[1]

ARTHUR G. STEINBERG AND BRENDA DAWN GILES

Department of Biology, Western Reserve University

INTRODUCTION

WICKERHAM (1956) demonstrated that the heterothallic yeast *Hansenula wingei* is composed of two mating types which agglutinate when mixed. The agglutination is more marked in heat-killed cells than in living cells (Brock, 1958a). It appears to be due to the presence of a protein on the surface of one mating type (called 21) and of a polysaccharide on the surface of the other (called 5) (Brock, 1958b and 1959a and b). The agglutination may be inhibited by various agents, among them being pooled human serum globulin (Brock, personal communication). This observation led us to examine individual human sera to see if they differed in their ability to inhibit the agglutination of the mating types of *H. wingei*, and if so, to see whether the difference is genetically determined. Our first series of ten tests showed that human sera do vary in their capacity to inhibit the agglutination of heat-killed cells of strains 5 and 21 of *H. wingei*. Accordingly, population and family studies were undertaken.

This paper is a report of studies of samples of serum from adult whites and Negroes and from white and Negro families. All samples, unless otherwise indicated, were tested within two weeks of collection, the majority within one week.

MATERIALS AND METHODS

Heat-killed cells of strains 5 and 21 of *H. wingei* were kindly supplied by Dr. T. D. Brock. The cells may be stored, at 4° C., as heavy suspensions in distilled water, up to six weeks. After this time they may show slight auto-agglutination. The suspensions were diluted with normal saline to an optical density of 0.2 at 512 mu ($\simeq$35,000 cells/ml.) in a Klett-Summerson colorimeter for use in the agglutination tests. All tests were done at room temperature in 10 x 75 mm. test tubes, as follows:

Controls: A. Agglutination: 0.1 ml. of saline, 0.1 ml. of strain 21, and 0.1 ml. of strain 5 were mixed in that order, although the order of addition does not seem to matter. The mixture was shaken by hand two or three times. Agglutination, readily visible with the naked eye, occurred in less than thirty seconds.

B. No agglutination: (a) 0.1 ml. of saline plus 0.1 ml. of strain 5 (b) 0.1 ml. of saline plus 0.1 ml. of strain 21.

Received August 20, 1959.

[1] Supported in part by research grant A-1822 from the National Institute of Arthritis and Metabolic Diseases of the National Institutes of Health.

Tests: 0.1 ml. of the serum to be tested, 0.1 ml. of strain 21, and 0.1 ml. of strain 5 were added in that order. The mixture was shaken by hand two or three times and allowed to stand for one hour at room temperature.

In no case in which serum had been added have we observed agglutination before five minutes. After many trials we have found that uniform results are obtained if the tubes are allowed to stand for one hour at room temperature and then read under a dissecting microscope at 9×. The cells are considered to have agglutinated if agglutinates of several cells are observed; otherwise, agglutination is considered to have been inhibited. The agglutinates varied in size, depending upon the individual serum used, from those readily visible with the naked eye (rarely) to those which were composed of fewer than ten cells (usually).

All mixtures could be caused to agglutinate strongly by centrifuging at 1,000 RPM for one minute. The inhibition, therefore, is not complete; similarly, the failure to inhibit is not complete since all sera cause a reduction in the avidity and the intensity of the agglutination. Nevertheless, the results are remarkable reproducible; many sera have each been tested at least twice, some as many as fifteen times, with identical results each time.

Samples of serum from white and Negro patients were obtained from the clinical hematology laboratory of the Out-Patient Department of Lakeside Hospital to determine the frequency of the inhibiting factor in these two populations. Samples of serum were obtained from members of families whose blood groups were being determined in the course of other genetic studies being undertaken in this laboratory.

Population data: Sera from 209 Negroes were tested and 99 (47.4%) were positive (i. e. inhibited agglutination); similarly, sera from 144 whites were tested and 65 (45.1%) were positive (Table 4, to be referred to again later). These frequencies are not significantly different ($\chi^2_{(1)} = 0.170$, $P > .5$). For the purposes of this report they will be considered to be the same. The frequency of positive sera in the combined sample is 46.5 per cent (164/353). Although not shown in the table, the frequency of inhibition is independent of sex.

Family data: Sera of parents and offspring in 55 families were tested. The data are shown in Tables 1, 2, and Appendix Table A. From the data in Table 1 it may be concluded that mating with regard to the character is random. The data in

TABLE 1. DISTRIBUTION OF MATINGS ACCORDING TO THE PHENOTYPE OF THE PARENTS

Female	Male		Number	
			Observed	Expected*
+	+		11	11.87
+	−		13	13.68
−	+		17	13.68
−	−		14	15.77
		Total	55	55.00

+ = inhibition.

$\chi^2_{(3)} = 1.102$ $P > .5$.

* Computed on assumption that frequency of non-inhibition is .5354 and that mating is random.

TABLE 2. DISTRIBUTION OF OFFSPRING BY TYPE OF MATING

Parents with + Serum	No. of Families	Offspring				
		Total	+		−	
			No.	%	No.	%
2	11	32	16	50.0	16	50.0
1	30	92	22	24.9	70	76.1
0	14	40	—		40	100.0
Total	55	164	38	23.2	126	76.8

+ = inhibition.

TABLE 3. RELATION BETWEEN AGE OF OFFSPRING AND PRESENCE OF INHIBITING FACTOR

One parent has factor				Both parents have factor			
Age (in years)	+	−	Total	Age (in years)	+	−	Total
0–9	17	46	63	0–9	8	7	15
10–	5	24	29	10–	8	9	17
Total	22	70	92	Total	16	16	32

+ = inhibition.

Table 2 indicate that the sera of offspring do not inhibit the agglutination of strains 5 and 21 of *H. wingei* unless the serum of at least one parent does. It is unlikely that this is a chance observation because the probability that the serum of none of the 40 children, from parents whose serum fails to inhibit the agglutination, would inhibit is approximately .0003 if the probability of failure to inhibit is taken as .768 (Table 2). The data show also that the frequency of inhibiting serum is probably greater when both parents inhibit than when only one parent inhibits ($P < .01$). A reasonable assumption is that the presence of the inhibiting factor is due to a dominant gene. Routine analysis of the detailed family data (Appendix Table A) indicates that the frequency of inhibitors among the offspring is too low in matings in which one or both parents inhibit to be in agreement with the hypothesis.

The frequency of inhibitors among the parents is .473. This is similar to that of the population sample, .465 (Table 4). The frequency among the children, however, is only .232 (Table 2), and this is significantly less than the frequency among the parents or the adult population ($P < .001$). Such data suggest that the presence of the inhibiting factor is age dependent. The data are too few to permit a detailed analysis, but separation into groups of children less than ten years old and those ten or older failed to reveal any suggestive differences (Table 3). Further comment will be reserved for a later section of the paper.

NATURE OF THE INHIBITING SUBSTANCE

The effect of storage: The data from samples tested at less than two weeks after collection and the data from samples tested after at least four weeks of storage at −20° C. were as shown in Table 4. (Some were samples previously tested but most were not so tested.)

TABLE 4. FREQUENCY OF INHIBITION OF AGGLUTINATION OF HEAT-KILLED CELLS OF STRAINS 5 AND 21 OF *Hansenula wingei* BY SERUM LESS THAN TWO WEEKS OLD (FRESH) AND SERUM MORE THAN FOUR WEEKS OLD (STORED)

Source of Serum	"Fresh" Serum			Stored Serum		
	Total	+		Total	+	
		No.	%		No.	%
Negroes	209	99	47.4	137	10	7.3
Whites	144	65	45.1	148	37	25.0
Total	353	164	46.5	—	—	—

+ = inhibition.

Comparison of the data for sera stored less than two weeks with those for sera stored more than four weeks shows for both races that there has been a significant reduction in the frequency of inhibition ($P < .001$ in each instance). As stated earlier, the frequencies of inhibition for 'fresh' sera from whites and from Negroes are not significantly different ($\chi^2_{(1)} = .170$, $P > .5$). A similar comparison for sera stored more than four weeks shows a significantly greater frequency of inhibition in sera from whites ($\chi^2_{(1)} = 16.203$, $P < .001$); hence, the loss of activity appears to be slower for sera from whites than for sera from Negroes.

The reduced frequency of the presence of the inhibiting factor among children suggests that the inhibition involves an antibody-antigen reaction. This was tested by absorbing two parts of active serum with one part of packed cells of strain 5 or strain 21, and by strain 5 followed by strain 21, at 0° C. for five minutes. In all three cases the absorbed serum failed to inhibit the agglutination of the yeast cells. Similarly absorption with brewer's yeast or baker's yeast removed the inhibiting activity.

Studies in progress with Dr. Irwin Lepow on the serological nature of the inhibiting factor in which direct tests of complement were made indicate a requirement for the four components of complement and for a factor resembling antibody in its ability to be absorbed from a positive serum by treatment with yeast at zero degrees. The results of these studies will be reported at a later date.

DISCUSSION

The population data indicate that about 47 per cent of adult whites and Negroes have a factor or factors in their serum which inhibits the agglutination of heat-killed cells of strains 5 and 21 of *H. wingei*. The family data suggest that the presence of this factor may be due to a dominant gene, but the reduced frequency of the factor among the children indicates that this simple hypothesis is not adequate to explain the data. The data could be explained if it is assumed that the inhibition is due to an antibody and that the ability to form this antibody is genetically determined. The increased frequency of antibody among sera from older individuals would then be due to their greater exposure to the risk of contact with the antigen. The antigen is present in brewer's yeast, baker's yeast and probably in other yeasts; hence, it is likely that all individuals are regularly exposed to it.

SUMMARY

A serum factor detected by its ability to inhibit the agglutination of heat-killed strains 5 and 21 of *H. wingei* is described.

The factor occurs with equal frequency in white and Negro adults (46.5% of 353). It is less frequent in children (23.2% of 164).

Experiments in progress suggest that each of the four components of complement and a factor resembling antibody are necessary for the inhibition to occur.

The ability to form the antibody appears to be dependent upon a dominant gene.

REFERENCES

BROCK, T. D. 1958a. Mating reaction in the yeast *Hansenula wingei*. *J. Bact.* 75: 697–701.

BROCK, T. C. 1958b. Protein as a specific cell surface component in the mating reaction of *Hansenula wingei*. *J. Bact.* 76: 334–335.

BROCK, T. D. 1959a. A biochemical basis of mating in yeast. *Science.* 129: 960–961.

BROCK, T. D. 1959b. Mating reaction in *Hansenula wingei*. Relation of cell surface properties to agglutination. *J. Bact.* 78: 59–68.

WICKERHAM, L. J. 1956. Influence of agglutination on zygote formation in *Hansenula wingei*, a new species of yeast. *C. rend. Laborat. Carlsberg, ser. physiol.* 26: 423–443.

APPENDIX TABLE A. DISTRIBUTION AMONG THE CHILDREN OF SERA WHICH INHIBIT AND WHICH FAIL TO INHIBIT THE AGGLUTINATION OF STRAINS 5 AND 21 OF H. WINGEI

Both Parents' Sera Inhibit			One Parent's Serum Inhibits			Neither Parent's Serum Inhibits		
s	r*	n_{sr}	s	r	n_{sr}	s	r	n_{sr}
1	0	1	1	0	2	2	2	10
2	0	1		1	1	3	3	1
	1	2	2	0	2	5	5	2
3	1	2		1	3	7	7	1
	3	3		2	4			—
4	1	1	3	1	1			14
6	2	1		2	4			
		—		3	4			
		11	4	3	3			
				4	1			
			5	3	1			
				5	1			
			6	5	2			
				6	1			
					—			
					30			

s = sibship size.

r = number who fail to inhibit.

n_{sr} = number of sibships of size s in which r fail to inhibit.

BOOK REVIEWS

Evolution, Marxian Biology and the Social Scene

By Conway Zirkle. University of Pennsylvania Press, 527 pp. $7.50.

The "Genetics Controversy" in the Soviet Union has attracted wide interest since in the early Thirties it first came to the attention of scientists in other countries. The first reactions of geneticists outside of the U. S. S. R. revealed attitudes of incredulity and incomprehension. Was it really true that there was a responsible body of opinion that rejected on whatever grounds the great body of experimental evidence on the mechanism of heredity which had culminated in 1926 in Morgan's "Theory of the Gene"? When the published records of the final debates of 1948 left no doubt that Mendelian genetics had been outlawed by a political decision the great question became: *why*? Some geneticists continued to regard the whole matter as a temporary aberration which must perforce blow over as the political leaders became aware of the deception which had been practised upon them by persons like Lysenko who were ignorant of (or ignored) the results and methods of modern biology. Others saw in the suppression of genetics, the overt expression of Communist ideology and tactics traceable directly to the founders of the movement—Marx and Engels. Persons with this view took the matter seriously and did not consider wasted, as did many in the first category, the great amount of time and the vast number of printed words devoted to an attempt to understand the causes of the dramatic events of 1948.

The author of the book under review belongs in the latter category. As a student of the history of biological ideas he was well prepared, in the academic sense, to view the changes in Soviet Biology in perspective, especially since he had written in 1946 the best early history of the idea of the inheritance of acquired characters, an idea which ostensibly occupied a focal position in the formation of Communist policy vis a vis genetics. In evaluating his new book we have to ask the two essential questions: what has he tried to do, and how well has he done it?

In the first place, the book is written for a general audience, hence includes a history of the theory of evolution, of modern genetics, and of what the author calls, without ever defining it, Marxian Biology. The animating purpose however is not exposition and analysis but persuasion. The view to which the author would persuade the reader is that Marxist ideas have infected and now tend to subvert not only biology (overtly subverted in the U. S. S. R., covertly subverted elsewhere) but sociology and literature in the U. S. A. and Great Britain. A large part of the book is in fact concerned with documenting this claim of Marxian infiltration into the intellectual life of our times.

I pass over the extensive citations from writers on sociology, from Tennyson, Jack London and Shaw (infected Marxists, all) and choose as illustrative examples of the author's purposes and conclusions, his account of the development of eugenics and human genetics as influenced by Marxism. The slow development of human genetics is noted. This is in part due to the reluctance of geneticists to apply their scientific reasoning to their own species, and it is more than hinted (p. 495) that the lack of zeal of some of them is because they are leaning in the Marxist direction. But there is some progress in spite of them: "The true importance of human genetics is slowly being appreciated" (p. 271). To this is appended a footnote: "The Society of Human Genetics was organized in 1948. It publishes a periodical, *Human Genetics*." (The erroneous references in this footnote are typical of the light-hearted way in which both proper names and scientific terms are treated throughout the book:

A. A. Fisher, Timiriozeff, pliotropic, Reketsia, "such a series of mutant genes is called a multiple allele" on p. 209, and many others.)

Eugenics occupies a prominent place because "Eugenics may serve as a very elegant indicator of tacit and unsuspected attitudes among the biologists themselves" (p. 142). Marxists oppose eugenical doctrines because Marx and Engels refused to accept Malthus' views; hence biologists who express skepticism concerning the application of Malthus' ideas to human populations or to eugenics as promulgated by Galton may be suspected of being crypto-Marxists. Professor Zirkle does not want the constructive contributions of the early proponents of eugenics to be lost sight of, such as their evaluation of "the disgenic (sic!) effects of the prevailing differential birth rate . . ." (p. 143). He is even willing to accept the early conclusion that "the most easily recognized type of feeble-mindedness, however, seems to be due to a single, deleterious, recessive mutation, whose expression is complicated by an unknown number of modifying genes" (p. 492). His authority for this is a paper by E. M. East published in 1931. He has moreover computed (p. 272) the rate of elimination of this gene as an indication of what negative eugenics can do and follows this by pointing out that negative eugenics are anathema to Marxians. To point out the dubious nature of the assumptions underlying his computations would thus probably render his critic liable to classification as a Marxian. But he ends on a charitable note for his erring biological colleagues: "Perhaps, once they are able to identify the role of the Marxian line, they may be able to free themselves from its influence" (p. 498).

Finally, we may ask: Has Professor Zirkle proved his point? Has he shown that the "Marxian syndrome" has damaged the development of biology? Curiously enough, this essential claim remains entirely undocumented as far as non-Soviet biology is concerned. In respect to biology in the Soviet Union, the chief effect of what Zirkle calls Marxian Biology —namely the suppression of genetics—was the result of a political decision of the Central Committee of the U. S. S. R. Communist Party and the part that Marxism played in that decision is nowhere made clear. This reviewer would hazard the guess, that the chief historical importance of that decision will eventually be found to lie in the assertion of dominance of political over scientific authority. That is an issue which every society of the future must face—and it merits more sober study than it has received in this book.

L. C. DUNN
Department of Zoology
Columbia University

The Molecular Basis of Evolution

By CHRISTIAN B. ANFINSEN. New York: John Wiley and Sons, 1959, xiii + 228 pp. $7.00.

BOOK REVIEWS that quote extensively from the author's preface sometimes arouse the suspicion that the reviewer did not bother to read all (or any!) of the book itself. In the present case, however, this suspicion seems worth risking, since some of the prefatory remarks seem to offer important clues to the genesis and nature of the book. According to the author, this is "a highly personal volume expressing the experimental and philosophical outlook which has resulted from a process of self-education in an unfamiliar area of science. As with many biochemists, pure biology, including genetics, has not been a major exposure in my education, and the process of learning something about these subjects has been both a revelation and a struggle." These lines also serve to give a sample of the flavor of the author's style.

Judging by the evidence of this volume, the attempted self education of the author, who is Chief of the Laboratory of Cellular Physiology and Metabolism at the National Heart Institute, has been successful. He describes in clear and at times fascinating fashion modern information concerning such topics as the chemical nature of the genetic material, the substructure of genes, protein structure, protein synthesis, and the relation of the structure of proteins to their biological activity. This reviewer learned much that was new to him, and failed to note any errors in what was not new; he will thus have to forego the usual triumphant list of slips and misprints.

The title of the book, be it noted, is exact. It deals with the molecular *basis* of evolution, not with evolution itself. Readers taking up the volume with the expectation of finding a detailed discussion of evolutionary mechanisms and developments will be disappointed.

Since the book is personal and not obviously designed to fit neatly into any particular niche, it is hard to criticise it. Any deficiencies may well be quite deliberate. It might be suggested, however, that the discussions of evolutionary principles and classical genetics would hardly be adequate for the biochemical reader who had not studied these subjects before. Supplemented by some of the suggested reading, however, they might answer.

On the whole this is an attractively written and produced book which can be heartily recommended. It has a brief index.

WILLIAM C. BOYD
Boston University
School of Medicine

L'Hérédité en Ophtalmologie

By JULES FRANCOIS, Professor of Ophthalmology at the University of Gand, Belgium. Paris: Masson and Cie, 1958, pp. 867. 8000 fr.

DR. FRANCOIS, whose interest and adequate background in ophthalmic genetics has heretofore been manifested in his profuse writings pertaining to hereditary ocular disease, has divided his subject material into four sections: Section I—*General Genetics* (having particular reference to human genetics); Section II—*Ophthalmic Genetics*; Section III—*Hereditary Ocular Diseases*; Section IV—*Systemic Diseases Having Associated Ocular Manifestations.*

Section I, *General Genetics* is further subdivided into thirteen chapters which concisely but scantily cover the principles of general genetics. Even intermediate inheritance, nondisjunction of chromosomes, mutation and biometry are introduced to the reader.

Section II, *Ophthalmic Genetics* is divided into four chapters and emphasis is placed upon certain interesting peculiarities of hereditary human ocular diseases such as: (1) clinical variability of certain genetic ocular entities and (2) the importance (correctly) of recognition (diagnosis) of the "carriers" of dominant, recessive, and incompletely penetrable genes which occasionally produce recognizable ocular aberration. In a most admirable manner the author reveals that a considerable amount of hereditary ocular disease can be successfully treated prophylactically, surgically, and medically. It is regrettable (personal prejudice) that in Chapter 4, Section II, Dr. Francois does not adequately emphasize the need for careful evaluation of each individual and particularly that specific individual's pedigree when undertaking to give genetic prognostication. The author, it is true, gives a most complete list of ocular diseases and specifically designates their mode of inheritance, but fails to stress adequately (at that point in the book) the basic need to individualize each *patient's particular pedigree* carefully. This is most important since peculiarly in human ocular genetics the same clinical entity apparently may exhibit variations in modes of inheritance in different

families. Knowledge, therefore, of the usual mode of inheritance of a trait is of no value—in fact, may even be dangerous—in prognosticating how a gene will behave in a given individual's case. ***Each individual's pedigree must be taken and studied before any ocular genetic prognostication is even undertaken!*** In fact, if this is not done and false data given to an individual it is conceivably possible that the physician could be held legally responsible.

The inclusion of Cuendet's (1957) genetic risk table (pp. 153–155) for certain ocular hereditary diseases is a valuable adjunct for any ophthalmologist called upon to voice genetic prognostication.

Section III, *Hereditary Ocular Diseases*, covers a myriad of hereditary ocular diseases, presented in 18 chapters. The material is well illustrated, pedigrees (wherein possible to check) are accurate and, contrary to most texts, the traits are adequately described. Chapter 2 of this particular section deserves careful study by every ophthalmologist since adequate emphasis is made of the role of the gene in the causation of refractive errors. Chapter 3 is devoted to a sensible analysis of the genetic risk factors in the transmission of that dread ocular tumor of childhood—*retinoblastoma*. This knowledge is necessary to any practicing oculist.

Section IV, *Systemic Diseases Having Associated Ocular Manifestation*, contains eight chapters which discuss the ocular manifestations of genetic systemic diseases. The entities reviewed include a vast parade; to mention a few major divisions: (1) disorders of protein, carbohydrate and lipid metabolism, (2) diseases of the central nervous system, (3) diseases of the skeletal system, (4) dermatological diseases, and (5) diseases of the hematopoietic system. This section also includes an excellent chapter on *syndromes*—worth the price of the book! While the coverage of these syndromes is not complete (Who can keep up with this vastly expanding literature?), the reader should be stimulated to further study, a true test of the worth of any literary effort. Section IV has one particular weakness—undoubtedly unavoidable—and that is the apparent assumption by the author that the reader is thoroughly familiar with the systemic entities being covered. Emphasis is correctly placed on the ocular manifestations but ophthalmologists in general need additional information by which to diagnose the systemic disease.

Rumors to the effect that this book will soon be published in English are rife. If true, there should be no excuse for any conscientious ophthalmologist not having access to this fine collection of ocular genetic information.

HAROLD F. FALLS
University Hospital
Ann Arbor, Michigan

Cancer in Families: A Study of the Relatives of 200 Breast Cancer Probands

By DOUGLAS P. MURPHY AND HELEN ABBEY. Published for the Commonwealth Fund. Cambridge, Mass., Harvard University Press, 1959, x + 76 pp. $2.50.

THIS BOOK is an important addition to the literature on the genetics of human breast cancer since it provides a different study population and study design. The methods of investigation were essentially the same as those used for the study of uterine cancer published by Murphy in 1952.

The breast cancer probands were white women who had been treated for breast cancer (verified histologically) in any of 28 hospitals in Philadelphia and surrounding towns within the five-year period preceding the start of the study in 1949. The study was limited to pa-

tients who were living and thus available for interview. This restriction was made to permit comparison between the present data on breast cancer and those from the uterine cancer study, although no such comparisons are presented.

The control probands were white women who had registered at the University dental clinic. About half of the control probands for the uterine cancer study were selected from the same dental clinic. Fortunately, the former use of volunteers as control probands was not continued in the present design.

The authors present evidence to suggest that the omission of dead potential cancer probands did not significantly bias the results. Other reasons for loss might occasion more concern, however. If we exclude the potential probands who were dead or too ill to be interviewed, there were 364 breast cancer patients of which 200 (54.9%) were enlisted, while out of 410 living control probands, 198 (48.3%) were enlisted. These losses could represent a selective bias for socio-economic variables such as employment or frequent changes in residence. The fact that the rate of loss was similar in the two groups provides no assurance that any possible bias has been cancelled.

The control group was not deliberately selected to match the cancer group, but the degree of comparison on several variables seems reasonably adequate. The authors point out that comparability between the cancer and control *relatives* is more important than that between the two groups of *probands*. No data are presented concerning the comparability of relatives (except for total age distributions).

The information about relatives was recorded carefully, and reported operations were checked with hospital records and physicians. A "large proportion" of the deaths were verified by death certificates. The study is unique in having two phases of interviewing, the first inquiring about female relatives, and the second asking about male relatives and checking the original information. The small number of close female relatives added by this recheck indicates that the first records were essentially complete.

The presentation of the data would have been improved materially by separating the various categories of relationship to the probands. In almost all of the tables on cancer frequency the relatives are grouped into four categories: the female relatives in the cancer and in the control group, and the male relatives in these groups. Aunts and uncles by marriage should have been considered separately, as one type of control. When considering a condition as common as cancer, the cancer frequency among cousins would be expected to approach that of the population from which the families were drawn. The categories with principal genetic significance would be first the sibs, then the parents, and finally the aunts and uncles (brothers and sisters of the parents). The inclusion of the other groups increases the size of the sample at the expense of genetic meaningfulness.

As compared with some other studies the observed cancer frequencies among relatives are low. These lower rates may have resulted in part from the decision to limit probands to those aged 40 to 65 years. The exclusion of older probands (both cancer and control) would make the mean age of sibs (and others of the same generation) lower, and would thus reduce the exposure to the risk of cancer.

The analysis of the data showed no differences of statistical significance between the close relatives of the cancer probands and those of the control probands with respect to the frequency of breast cancer or cancer of other sites. This finding differs from the results of the majority of previous studies which showed an increased incidence of breast cancer in the cancer group. Although this is a significant point, careful comparison with the other studies is not possible since the age distribution of relatives cannot be compared with that in other studies.

The best indicators of any genetic factors in human breast cancer (whatever the mode

of inheritance may be) are the sisters of the probands. Among the 372 cancer sisters only 2 (0.5%) had breast cancer, while among the 379 control sisters there were 6 (1.6%) with breast cancer. It is of some interest that in the earlier study of uterine cancer (with approximately the same number of sisters) there were 4 cancer sisters and 2 control sisters with breast cancer. There is no reason to expect that the breast cancer rate among all sisters of breast cancer patients is really less than that among sisters of uterine cancer patients or sisters of suitable control probands. It would seem more probable that the effective sample size in the present study is so small that the proportion of cancer sisters and of control sisters having breast cancer is subject to considerable variability due to random fluctuations. We can agree with the authors' general conclusion which is stated fairly and conservatively: "If a familial tendency to develop cancer did exist, it was not large enough to be detected in a study of this size."

V. Elving Anderson
The Dight Institute
University of Minnesota

Les Méthodes en Génétique Générale et en Génétique Humaine

By R. Huron and J. Ruffié, Paris: Masson et Cie., 1959, viii + 556 pp. 8, 200 fr.

In a volume intended for those wishing an appreciation of the mathematical methods of genetics, the problem arises of how much basic genetics the reader can be assumed to know. Huron and Ruffié have chosen to assume that the reader knows none, and have explained the principles of genetics in some detail, taking a quarter of the volume to do so. The second part of the book opens with a short discussion of the application of genetic principles to man. After this, the book gets to the mathematical techniques to which it is basically devoted.

The volume gives all indications of having been put together inadequately and of never having been checked except for French grammar. The material on formal genetics is largely in general terms, with a few errors such as the definition of *dominant lethals* as genes giving visible effects in heterozygotes and lethality in homozygotes. A large proportion of the examples in formal genetics are drawn from the work on *Drosophila* and the misleading and erroneous statements concerning this animal are numerous. Human genetics has not fared as badly, though included among other odd statements one can read that females homozygous for hemophilia do not survive to birth according to data from "a great number of crosses" between carrier females and hemophilic men. At the Rh locus, *D* and *d* are said to be codominants.

The mathematical section does not contain any highlights. The authors adopt some curious views on statistics. For example, they discuss a paternity suit with an AB mother and a B child. With the gene frequencies of France, the true father would be of type AB only three per cent of the time, and it is implied that with so small a frequency a geneticist should testify in court to the unlikelihood of guilt of an accused AB man. The reader who fails to see the error of this example is strongly cautioned to avoid the other examples in this book.

The two-page bibliography is remarkable in that it may require more editing than any other unit of comparable size in modern scientific printing, probably requiring more than two changes per line. For example, seven sets of papers originally from this Journal are quoted. For these seven instances, the name of this Journal appears in five forms, none of which is acceptable.

It is regrettable that so unpleasant a review has to be written about any book, but the prospective reader deserves fair warning, and he must admit that he has it here.

H. M. SLATIS
Argonne National Laboratory
Lemont, Illinois

La Dystrophie Myotonique (Steinert) et la Myotonie Congenitale (Thomsen) en Suisse. Etude Clinique, Genetique et demographique.

By D. KLEIN. Supplement to: Journal de genetique humain, *1*, 1958, 328 pp., 289 figures, 14 tables. (In French, with a short English and German summary).

IN HIS MONOGRAPH on clinical, genetic and demographic aspects of myotonia dystrophica and myotonia congenita in Switzerland, Dr. Klein presents the results of eleven years of a carefully planned investigation of all known cases of both diseases in Switzerland.

The material has been collected by following the few previously described Swiss myotonic families, but mostly by enquiries of all Swiss neurologists and ophthalmologista and in all hospitals and homes for the mentally ill. Thus data were collected on 319 cases of myotonia dystrophica and on 19 cases of myotonia congenita, of whom 299 patients with myotonia dystrophica and 18 with myotonia congenita were examined by the author. In addition are listed a further 19 cases of myotonia dystrophica and 6 of myotonia congenita recorded after the completion of the manuscript.

The reviewer was able to appreciate the extensive scope of Dr. Klein's investigation, and considers the above numbers to cover practically all cases of both diseases in Switzerland. Dr. Klein finds thus 4.9 patients with myotonia dystrophica per 100,000 inhabitants and estimates the mutation rate at about 1.6×10^{-5}; in myotonia congenita he finds 0.38 cases per 100,000 inhabitants. These numbers should be slightly higher, as the 19 and 6 cases respectively, diagnosed after 1956 and listed after the completion of the manuscript, were not included in the computation. The distribution of the patients shows a certain concentration of cases in some relatively isolated groups of population.

On each patient examined there is an extensive family study, a case history and a concise but complete clinical description usually accompanied by a characteristic and well reproduced photograph.

Due to the far greater number of cases and the much more complicated clinical picture, most of the monograph is devoted to myotonia dystrophica. Myotonia and myatrophia, semeiotic particularities of this disease show a varying degree of expression and of incidence in different muscle groups. Ninety-eight per cent of patients suffered from early or late states of cataract. There was a diminution, difficult however to estimate, of sexual function in 70% of the male and in 47% of the female patients, with the resulting decreased fertility of 1.9 as compared to the 2.5 average of the Swiss population. Psychical modifications were present in 35% of the patients, of which 22% were diagnosed as oligophrenic and 13% as disorders of the affectivity or of the character. From the 24 cases examined by ECG, 18 were found to be pathological. The case histories also reveal the incidence of other pathognomonic traits such as sex limited early baldness, ophthalmic pathology other than cataract, cardiac pathology, neuropathology and involvement of other systems and functions too numerous to be repeated here.

The average age of the onset of the disease, calculated for all patients, is 25.4 years with a prevalence in the 30 to 34 year age group. However, when divided into generations, the average age of onset is 50.5 years for the parental and 27.4 for the filial generations which,

the author presumes, represents the phenomenon of anticipation (anteposition) of the disease in the course of successive generations. He is fully aware of objections raised to this hypothesis by Penrose. However he refers (pages 46 and 47) to anticipation observed by Goldschmidt in *Drosophila*, and to similar reports of other authors on myotonia dystrophica. The author confirms the dominant autosomal mode of inheritance of the disease and, in spite of the equal sex-ratio among patients, the masculine sex predominates in transmitting the disease.

The cases of myotonia congenita described confirm the classical picture of this disease; the differential diagnosis with myotonia dystrophica is thoroughly discussed.

On the whole the monograph of Dr. Klein brings not only new and valuable genetical data but is of great value also as a source of reference for diagnosis of the sometimes very atypical forms of myotonia dystrophica.

JAN K. MOOR-JANKOWSKI
National Institutes of Health
Special Research Fellow
Blood Group Research Unit,
The Lister Institute, London, S.W.1

LETTER TO THE EDITOR

A Law for Human Genetics

To the Editor
Dear Sir:

It has been apparent for a long time that human genetics has important social applications. Counseling in human genetics fills a practical need for those able to obtain it. Clients have come in person to the Dight Institute from practically every state for this service. Adequate facilities to meet the demand should be present in every state, and eventually they will be established.

In Minnesota the Genetics Research Program at the Rochester State Hospital has developed a state-wide interest in genetic research and counseling far beyond what might be expected considering the minor importance of the disease investigated, Huntington's chorea, in the total health picture. Only a momentary descent from our academic towers is necessary to discover the popular interest in the subject of heredity and its social implications, and the desire for some sort of action in the field for the benefit of everyone.

Laws concerning marriage, divorce, state guardianship, and other topics which may be related to genetic changes in the population are formulated at the state and county levels. This is also true for regulations of the more conventional sort concerned with water pollution, epidemics and health problems of every description. Each state has its Board of Health which is technically responsible for the genetic health of the population as well as environmental factors such as parasites and sanitation. The various Boards of Health have been highly successful in their struggles with the environment but the gene has eluded them completely.

Dr. John S. Pearson, chief psychologist at the Mayo Clinic and president of the Minnesota Human Genetics League, led a successful movement intended to interest the State Board of Health in the usefulness of a human genetics unit in the State Board of Health and to obtain authorization in the form of a law for initiating such a unit. All state agencies wish to have their powers and duties clearly defined by acts of the legislature and such action is prerequisite to success. Wonderful cooperation from everyone approached was obtained and a bill authorizing the State Board of Health to conduct a Program for Study of Human Genetic Problems, and establishing an advisory committee was passed unanimously by the Minnesota House and Senate. The bill became a law on the twenty fourth of April, 1959 with the signature of Governor Freeman.

The new law hardly elevates the gene to the same exalted status as *E. coli* but it is official recognition that man has genes and that something can be learned about them and that presumably the information can be utilized to advantage.

SHELDON C. REED
Dight Institute for Human Genetics
University of Minnesota

CORRECTION

Correction to the article by CURT STERN, The Chromosomes of Man, Reprinted in Volume 11, Number 2, Part 2 from the Journal of Medical Education.

The idiogram as figured consists of only 45 instead of 46 chromosomes. The 46th was correctly included in the engraver's proof but cut off at the right of the chart when the final copy was arranged by the printer!

INDEX

ABO blood groups
  and sex ratio, 274
amylase activity of saliva, 265
antigens
  Di^a, 17
  P^k, 28
ASHLEY, D. J. B., 305
Bayes' theorum
  and linkage investigation, 289, 298
BECK, S., 215
BENNETT, J. H., 169
β-aminoisobutyric aciduria
  biochemical genetics of, 257
blood groups
  ABO and sex ratio, 274
  Di^a, 17
  P, 26
  selection of MN, 237
books reviewed
  Bilateral Polycystic Disease of the Kidneys, 83
  Cancer in Families, 388
  Cytodifferentiation, 284
  La Dystrophie Myotonique (Steinert) et la Myotonie Congenitale (Thomsen) en Suisse. Etude Clinique, Genetique et demographique, 391
  The Evolution of Development, 80
  Evolution, Marxian Biology and the Social Scene, 385
  The Fertility of American Women, 203
  The Genetic Basis of Selection, 84
  Genetics, A Survey of the Principles of Heredity, 86
  Hemophilia B: Genetics, Hematology and Clinical Aspects, 205
  L'Hérédité en Ophtalmologie, 387
  Heredity Counseling, 207
  Heredity and Evolution in Human Populations, 203
  Inborn Errors of Metabolism, 279
  Looking at Chromosomes, 206
  Les Méthods en Génétique Générale et en Génétique Humaine, 390
  The Molecular Basis of Evolution, 386
  Perspectives in Marine Biology, 204
  The Population Ahead, 281
  Population Studies: Animal Ecology and Demography, 280
  Radiographic Atlas of Skeletal Development of the Hand and Wrist, 282
  Report of the United Nations Scientific Committee on the Effects of Atomic Radiation, 81
  Safe Handling of Radioisotopes, 206
  Über die Erblichkeit des normalen Elektroenzephalogramms, 90
  Variables Related to Human Breast Cancer, 88
  Zur Kenntnis des "Mongolismus", 283
caries, dental (see dental caries)
CHU, E. H. Y., 63
CHUNG, C. S., 237, 339, 360
COHEN, BERNICE H., 274
dental caries
  heritability in, 263
  pH, 265
  twins, 263
dermatoglyphics, Palmar, 252
DEVI, ILA, 35
Di^a, antigen
  inheritance of, 17
    independence of
      other blood group systems, 17
      private, 23
      public, 23
DONNELLY, A. J., 215
Duchenne
  muscular dystrophy, 351
dystonia musculorum deformans
  dominant cases, 194
  idiopathic, 188
  recessive cases, 197
ectopia lentis et pupillae dominant, 166
eugenic advice, 303
eye brows, 35
facioscapulohumeral muscular dystrophy, 354
FALEK, A., 52
GARTLER, S. M., 257
genetic entities
  in muscular dystrophy, 339
genetic tests
  incomplete ascertainment, 1
GILES, B. DAWN, 380
GILES, N. H., 63
GLASS, B., 274
GOODMAN, H. O., 263
GRAHAM, J. B., 311, 333
HACKEL, E., 263

handedness, 52
HEATH, G. G., 166
HSIA, D. Y. Y., 156
human chromosome complements, 63
HUNGERFORD, D. A., 215
Huntington's chorea
  fertility, 118
  fitness, 121
  in Michigan, 107
  mutation, 107, 126
  selection, 107
hypophosphatemia, 156
  mode of inheritance, 161, 319
  penetrance, 319
  with vitamin D-resistant rickets, 311
intersex, phenotypic
  chromosome constitution 215, 221, 228
  sex chromatin, 221
JENKINS, J. T., 188
KAEIBLING, R., 188
kidney disease, 333
KRAUS, M., 156
kuru
  possible genetic basis, 169
  sex and age distribution, 173
LAYRISSE, M., 17
letters to the editor
  Albinism: reply to Dr. Witkop, 91
  Fertility differential in two Lappish populations, 92
  A law for human genetics, 393
  Tyrosinase and albinism, 91
limb-girdle muscular dystrophy, 354
linkage investigations
  Bayes' theorem, 289
  sequential tests, 293
  statistical methods, 289
LUKE, J. E., 263
MATSON, G. A., 26
MCFALLS, V. W., 311
metaphase chromosomes, 228
microflora in dental caries, 263
MN blood groups
  selection of, 237
MORTON, N. E., 1, 237, 339, 360
muscular dystrophy, 360
  discrimination of genetic entities, 339
  Duchenne, 351
  facioscapulohumeral, 354
  limb-girdle, 354
mutation
  Huntington's chorea, 107, 126
NEEL, J. V., 107
NOADES, JEAN, 26
NOWELL, P. C., 215
ovarian pregnancies
  parthenogenetic, 305
P blood group system
  anti-P^k, 28
  "new" antigen, 26
Palmar dermatoglyphics
  genetics, 252
  sib pair correlation, 255
parthenogenetic ovarian pregnancies, 305
PASAMANICK, B., 188
pH and dental caries, 265
PONS, J., 252
RACE, R. R., 17, 26
REED, T. E., 107, 137
relative fitness
  of specific genetic traits, 137
RHODES, F. A., 169
rickets
  vitamin D-resistant, 156, 311
    mode of inheritance, 161, 319
    penetrance, 319
ROBSON, H. N., 169
ROSEN, S., 263
saliva amylase activity, 265
salivary components in dental caries, 263
SAMUELS, JUNE, 156
SANGER, RUTH, 17, 26
selection
  complete, 2
  Huntington's chorea, 107
  incomplete
    multiple, 2
    single, 2
    truncate, 2
  MN, 237
sequential tests applied to linkage, 294
serum factor
  effect on yeast, 380
sex ratio and ABO blood groups, 274
significance tests, 291
SMITH, C. A. B., 289
STEINBERG, A. G., 380
SWANSON, JANE, 26
twins
  dental caries in, 263
vitamin D-resistant rickets, 156, 311
  mode of inheritance, 161, 319
  penetrance, 319
WALLS, G. L., 166
WINTERS, R. W., 311
ZEMAN, W., 188